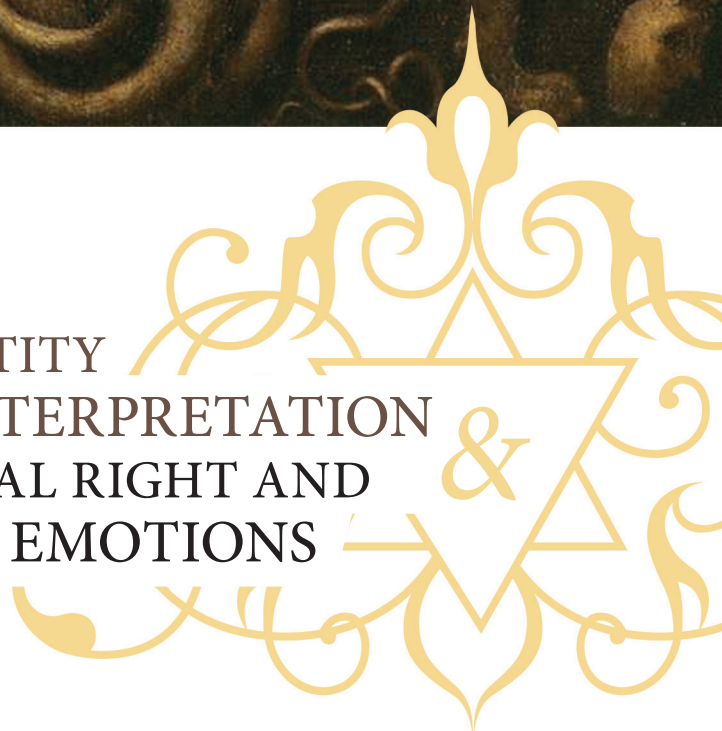




PERSONAL IDENTITY  
AND SELF-INTERPRETATION  
NATURAL RIGHT AND  
NATURAL EMOTIONS



Personal Identity and Self-Interpretation  
& Natural Right and Natural Emotions

Budapest Seminars in Early Modern Philosophy 2 & 3



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

Personal Identity  
and Self-Interpretation  
& Natural Right  
and Natural Emotions

Budapest Seminars in Early Modern Philosophy 2 & 3

Cover: The Portrait's Box by Ridolfo del Ghirlandaio (1506–1510; Uffizi Museum)

Photo: © Web Gallery of Art, 2011

© Authors, 2020

© Editors, 2020

ISBN 978-963-489-228-1

ISBN 978-963-489-229-8 (pdf)



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

---

Executive Publisher: the Dean of the Faculty  
of Humanities of Eötvös Loránd University

Layout: Ádám Bornemissza

Cover: Ildikó Csele Kmotrik

Print: Multiszolg Ltd



# Table of Contents

Introduction . . . . .	7
PERSONAL IDENTITY	
Judit Szalai: States of the “Person” in Descartes . . . . .	13
Przemysław Gut: Leibniz on Personal Identity . . . . .	22
Charles T. Wolfe: Diderot and Materialist Theories of Self . . . . .	37
Ákos Forczek: Apperception and Affinity: Kant on the Material Condition of the Identity of the “Psychological Person” . . . . .	53
SELF-KNOWLEDGE	
Bartosz Żukowski: Richard Burthogge’s Epistemology and the Problem of Self-Knowledge . . . . .	69
Peter West: Knowing Me, Knowing You: Berkeley on Self-Knowledge and Other Spirits . . . . .	84
NATURAL LAW	
Wilhelm Schmidt-Biggemann: The Space for Emotions in Natural Law . . . . .	107
Szilárd Tattay: Francisco Suárez and Modern Rationalist Natural Law Theories . . . . .	121
Gábor Boros: Natural Right and Worldly Emotions of Love: Descartes, Molière, Spinoza . . . . .	139
Heikki Haara: The Concept of Simple Esteem in <i>De jure naturae et gentium</i> . . .	153
József Simon: Shame, Common Wealth and Religion in the Thought of Miklós Bethlen (1642–1714) . . . . .	167

Mariangela Priarolo: Love and Order: Malebranche and the Feeling of Natural Law . . . . .	180
Paolo Santangelo: Chinese Cultures of Love: The “Cult of <i>Qing</i> ” and <i>Qing</i> ’s “Naturalness” . . . . .	198
LIST OF CONTRIBUTORS . . . . .	221

## Introduction

At least from the time when Descartes announced that the mind of the subject is better known to itself than the physical world, self-knowledge and personal identity played a central role in early modern philosophy. The focal questions in this period include whether the subject can be identified with the knowing mind, whether personal identity is grounded in the identity or some properties of that mind, or is in some manner based in physical reality, and whether self-knowledge is indeed superior to other types of knowledge. Some of the classical answers were given by prominent early modern thinkers, characteristically aiming at compatibility with the new mechanistic natural sciences. Other, lesser-known, but at times comparably promising solutions were offered by philosophers some of whose ideas we have yet to assimilate.

This volume is based on papers presented at the Second and Third Budapest Seminar in Early Modern Philosophy held at the Eötvös Loránd University of Budapest on 26–27 November 2017, and 8–9 October 2018, respectively. We do not aim at presenting a systematic survey of personal identity, self-knowledge and natural law in early modern philosophy. Instead, we sample texts by less visible authors and re-think better-known positions concerning affectivity and the related conceptual fields in the period. The papers published come from different traditions, and some of them are already in implicit or explicit dialogue with others. We hope that this collection will contribute to the growth of our knowledge about the role of personal identity, self-understanding and natural law in early modern philosophy, and generate a greater exchange of ideas in the broader early modern scholarship.

We would like to thank the National Research, Development and Innovation Office (Hungary) for generous funding in the framework of a research project on “Self-Interpretation, Emotions, Narrativity” (K120375), which allowed both the organizing of the conferences and the publication of the present volume. We would also like to thank the Student Union of the Faculty of Humanities at ELTE (ELTE BTK HÖK) and of ELTE



(EHÖK) for the funding they provided, both for the organization of the conferences and the publication of this volume. Finally, we would like to express our special thanks to Ákos Forczek, who has helped us throughout the entire process.

**Judit Szalai**'s paper revisits the Cartesian notion of the "person". For Descartes, a person is constituted by a mind conjoined with a body; we know about the mind-body unity via functions that are both mental and physical. How can Descartes spell out these functions in terms of actual states of the mind and the body, given the narrow metaphysical constraints established by his own philosophy? Szalai's paper traces the psychophysical processes involved in the passions in order to see how they can be made sense of in the light of Descartes' metaphysical and epistemological commitments.

Remaining in the Cartesian tradition, **Przemysław Gut**'s paper targets those interpretations of Leibniz's account of personal identity, according to which Leibniz failed to formulate a coherent theory. Gut argues that Leibniz held the continuity of both the substance and the psychological phenomena necessary for personal identity, because both are grounded in the continuity of the existence of a substantial principle. Gut then identifies this substantial principle with the soul, i.e. the Cartesian thinking "I" that is the ultimate subject of both personal identity and moral responsibility, which are in the focus of Leibniz's philosophy.

Leaving Cartesianism and the seventeenth century behind, **Charles T. Wolfe**, in his highly engaging paper, takes a look at the broad implications of materialism for early modern theories of the self. His essay accomplishes two things simultaneously. First, it presents Diderot's materialist theory of the self as constituted by an externalist metaphysical theory and a biological understanding of individuality. Second, it manages to place this early modern development in the much broader context of the history of western philosophy. He argues for the necessity of a revision of the traditional dichotomy between metaphysically sound but mechanistic materialism lacking a viable philosophy of mind and dualism with a sound philosophy of mind built on a shaky metaphysical theory.

The last paper in the personal identity section focuses on Immanuel Kant, the key figure of the transition from early modern philosophy to German Idealism. **Ákos Forczek** in his contribution links Kant's major critical writings with his less studied texts in his *Opus postumum*, and maintains that "transcendental affinity" and personal unity are closely related concepts, even though at first glance this is far from obvious. His main focus is the problem of what guarantees that the sensible manifold is suitable for experiencing it without violating the autonomy of the sensible. Forczek argues that after the failed attempt of the first *Critique*, and the question-begging answer of the third, the true answer, related to self-knowledge, can be found in the *Opus postumum*.

Opening the section on self-knowledge, **Bartosz Żukowski**'s paper examines the lesser-known Richard Burthogge's philosophy of mind. Burthogge was an English physician at the end of the seventeenth century, whose views on the mind were considered highly unusual in his time and therefore remained unappreciated both by his contemporaries and by posterity. Żukowski shows, however, that behind Burthogge's idiosyncratic language,

the epistemology of an interesting and innovative thinker can be reconstructed, someone who shared some affinities concerning self-knowledge with Leibniz and Kant.

In contrast with Burthogge's idealism, which has been largely ignored by early modern scholarship, that of George Berkeley has received much attention. **Peter West** takes a look at the classical problem of self-knowledge in Berkeley's idealism: if ideas in the human mind represent their objects due to their similarity to external objects, and spirits are of a different metaphysical kind than ideas, how is self-knowledge, i.e. knowledge of the spirit that is me, possible? West argues that, given the empiricist commitments of Berkeley, the answer has to be relevantly similar to the way in which knowledge of external objects is possible, i.e. as immediate knowledge. The available object of knowledge in the case of self-knowledge then has to be the effects of the spirit.

**Wilhelm Schmidt-Biggemann's** contribution opens the section on natural law, in which he traces the early modern debates on the role of emotions and natural law to their medieval roots. He shows that in Aquinas natural emotions play an important role in constituting natural right, which is denied by Hobbes's conception of positive law. Finally, Pufendorf rehabilitates the notion of natural emotions grounding natural right, but in a new, civil sphere. He claims that while Pufendorf's philosophy allows more civil freedom in the state than either that of Aquinas or Hobbes, his argument echoes the traditional Aristotelian notion and appreciation of playing games in Aquinas's philosophy.

The focus of **Szilárd Tattay's** paper is similarly the complex interplay between scholastic and early modern philosophy. He places the natural law theory of the Spanish Jesuit Francisco Suarez, widely influential throughout the early modern period, within the scholastic debates on voluntarist and rationalist theories of natural right. He argues, by reconstructing Suarez's views on the formal basis of natural law, that Suarez tries to find a middle ground between what he considered to be the extreme voluntarist and rationalist positions. He argues that by doing so, Suarez remained firmly on the grounds of ecclesiastic scholasticism, without being a forerunner of secularist theories of natural law.

**Gábor Boros** also takes a look at the roots of the early modern notion of natural right. Whereas Schmidt-Biggemann's contribution mainly focuses on the medieval roots, and Tattay's on the Early Modern Scholastics' influence on Pufendorf, Boros's emphasis is on the ancients, the Stoic predecessors of this concept. More precisely, he shows that the normativity entailed by different branches of early modern thinking on natural right is derived from the different interpretations of the term "nature" in different authors, or within different passages of works of the same author. He argues that this ambiguity is paralleled by the ambiguous ways in which early modern philosophy has conceived the notion of love.

**Heikki Haara's** contribution picks up the thread of Schmidt-Biggemann's paper and places Pufendorf's theory of the emotion of simple self-esteem, which has implications for both his theory of natural law and moral philosophy, in the context of contemporary theories of esteem, influenced by Kant. In his reconstruction, esteem signifies the moral value of a person in a way that is modelled on the notion of price denoting the economic value of objects. Haara argues that the comparison between recent theories of esteem based on

one's accomplishments as a productive citizen and that of Pufendorf can be relevant to our contemporary concerns.

**József Simon's** paper discusses a Hungarian student of Pufendorf, who also happened to be a leading politician of his time, Miklós Bethlen. Simon provides a case study of the complex ways in which the early modern cultural context and the philosophical interests of Bethlen intersected and determined the way in which his philosophical thoughts were shaped and produced. He argues that Bethlen's materialist understanding of speech acts embeds a Pufendorffian conception of the role of emotions in grounding natural law, which provides the framework of Bethlen's thinking about political philosophy, and specifically his own role in politics.

The author at the centre of **Mariangela Priarolo's** contribution lived roughly in the same period as Miklós Bethlen, yet their philosophical interests were quite dissimilar. Priarolo's paper investigates Malebranche's ethical theory and, specifically, what roles emotions play in grounding natural law. Priarolo's analysis tracks the changes of the role Malebranche assigned to pleasure in his moral theory and argues for the fundamental importance of emotions to his ethics. Her central claim is that the way in which Malebranche formulates his considered view can be regarded as a forerunner and inspiration for sentimental moral theories, such as Hutcheson's.

The final paper of this volume, **Paolo Santangelo's** study of the cult of *qing*, provides an overview of a theoretical tradition contemporary and parallel to the European early modern theories of emotion discussed in other papers. Santangelo sketches the development of Chinese theories and discussions of what could roughly be labelled "love". His overview of what can be considered "natural" or "genuine" concerning love and intimacy in this parallel tradition can shed light on both the idiosyncrasies and the universal features of theories of natural emotions in philosophy.

# PERSONAL IDENTITY



## States of the “Person” in Descartes

For Descartes, the “person” is constituted by a particular mind and a particular body, enjoying a special, if somewhat obscure, connection. Functionally, what holds the “person” together are shared functions of the two substances, such as perception, memory and emotion. How can mind and body share functions without sharing *states* as well, which Descartes’ ontology hardly seems to allow? This is the theme of the present paper.

### **How Do Physical and Mental States Relate to Each Other?**

What is the relation between a feeling caused by a physiological state and the physiological state itself? Are they of the same kind, of radically different kinds, or is there a qualified difference between the two? If anyone’s, Descartes’ conception would seem to be unbendingly dualist: the two states, mental and physical, are independently identifiable states of things of very different natures, and as such have nothing in common, being from our viewpoint arbitrarily coordinated by God.<sup>1</sup>

An argument against this presents itself on the basis of Descartes’ general views regarding causation. Applying the “causal reality principle”, it seems that if a movement in the body causes a change in the mind, there has to be a common element in both. On the causal reality principle, a cause has to have as much reality as its effect. “As much reality” can be interpreted in two different ways. At times it appears to be a global, quantitative notion, as if different beings had different levels of overall reality and whatever is on a higher

---

<sup>1</sup> The “natural institution” view lays particular emphasis on this arbitrariness; see Wilson.

level of overall reality could produce an effect on a lower level. At other times it appears to be qualitative and specific, meaning that a mode (of some substance) cannot be produced without that very mode being present in the causing substance to at least the same degree. This qualitative interpretation is supported by some statements by Descartes, including the following:

The fact that there is nothing in the effect which was not previously present in the cause, either in a similar or in a higher form, is a primary notion which is as clear as any that we have; it is just the same as the common notion “Nothing comes from nothing”. For if we admit that there is something in the effect that was not previously present in the cause, we shall also have to admit that this something was caused by nothing. And the reason why nothing cannot be the cause of a thing is simply that such a cause would not contain *the same features* as are found in the effect. (AT VII 135/CSM 2.97. e.m.)

If we accept the qualitative interpretation, then, when talking about the body, or a state of the body, causing a state of the mind, the body would have to have the same mode in some manner. This cannot mean that the body should be capable of having a mental mode, or that there should be something physical in the mind. What *can* it mean, though?

The causal reality principle is put in the idiom of modes and substances. The meaning of “mode” is quite inclusive, however: a mode is something the existence of which depends on another thing (the substance of which is a mode) and is susceptible to change. In *Notae*, an example Descartes gives of “things which by their very nature are susceptible of change” is “the fact that at present I am writing or not writing as the case may be” (AT VIII B 347/CSM 1.297.). In the French version of *Meditations*, Descartes talks about “shapes, movements and other modes or accidents of the body” (AT VII 78/CSM 2.54.). Thus, not only shapes, sizes, etc. are modes, but “my writing right now”, or the modes involved in it, are also modes of my body. So if, according to the causal reality principle, there is a mode that is common between mind and body in interaction, that mode could be some quality, or also a state or event. (For considerations spelled out below, I will opt for taking the shared mode to be an event.)

If something has to be in common between mind and body in order for the two to be able to enter into a causal relationship, what could that something be? There is some reason to think that such commonality could be the very changes occurring in instances of interaction.

That there is an event-like mode shared by mind and body is suggested by the following. At the beginning of *The Passions of the Soul* we read: “although an agent and patient are often quite different, an action and passion *must always be a single thing* which has these two names on account of the different subjects to which it may be related” (AT XI 328/CSM 1. 328. e.m.). Specifying with reference to mind and body: “we should recognize that what a passion in the soul is usually an action in the body” (ibid., e.m.).

These lines may be surprising if we read them with Descartes' dualism in mind. Modes are supposed to belong to one substance only, and there is no event or change without a substance. If a mode belongs to two substances at the same time, the question concerning what makes those substances capable of sharing a mode, and the impression that they again do not appear to be as independent as strict dualism would have it, arises.

Besides the passion of the mind and the action of the body being "the same thing", the following consideration also speaks against the two-event interpretation of psychophysical interaction. *Causes* are typically *things* (with certain properties) for Descartes, rather than events (God is "the efficient cause of things" [AT VIII A/CSM 1.202] or "the primary cause of motion" [AT VIII A/CSM 1.240]; my ideas of different things may be caused by myself or something outside myself). The causal reality principle also assumes that causes are substances (minds or bodies). Causation is one thing acting on another, which changes some mode or property of the thing acted on, rather than one event following the other. Descartes' model of causation is much closer to mediaeval conceptions than could be considered a precursor to the Humean one.<sup>2</sup> When two bodies collide, the property of motion gets imparted or removed, partially or wholly.<sup>3</sup> If we wish to formulate the causation involved in terms of events, there is *one* event, that of imparting or removing a property.

Descartes scholars may have been attracted to the constant conjunction view of causation for the reason that such a view would fit a dualistic conception in a neat manner. For, if Descartes had such a view, the two events involved in causation, on the face of it, would not be puzzling: one of the event, clearly separable and independently identifiable, would be a physical one, the other a mental one. On closer inspection, though, this solution has its problems. For one thing, if there were two events following upon each other, Descartes' claim that the passion of a thing is the action of another, would not make sense. Second, the pineal gland is supposed to mediate between the brain and the mind. If there is an independently identifiable physical event followed by, that is, causing an independently identifiable mental event, the pineal gland seems to have no role.

The "natural institution view" would incline us toward attributing a constant conjunction conception to Descartes. However, this view does not tell us anything about the causal process involved. Rather, what it asserts, with deliberate vagueness, is that, in the case of the passions, certain movements of the body are associated with certain modes or states of the mind. Furthermore, it is a theory about the *content* of the mental modes involved: they do not resemble the bodily movements; thus, from our point of view, they are arbitrary. However, the substance-mode ontology is more important here: whereas we might think of (having) states *as events*, they are *modes* of substances for Descartes. The idea of the coordination of physical and mental modes does not commit him to a two-event view of causation.

---

2 For some aspects of the relationship between Descartes' views on causation and those of certain mediaeval authors, see Schmalz 2008.

3 See the "transference model of causation", as attributed to Descartes: Clatterbaugh 2009.



While a single-event interpretation would sit better with much of what Descartes says, it seems to stretch, or go past, the boundaries of Descartes' ontology. "Events" are not a separate category in Cartesian ontology, which operates with substances and properties of substances. In the sense, whether a one-event view is closer to the Cartesian spirit than a two-event view, thinking of causation in Descartes in terms of events is somewhat stretching it.

However, while we may not want to have events as spatiotemporally identified items (Davidson-style) in Descartes, there is room for understanding events as objects' acquiring or losing properties. Two bodies colliding, for instance, seems to involve a double property change: one of the bodies loses "a quantity of motion equal to that which it imparts to the other body" (AT VIII A 65/CSM 1.242). We could ask how this differs from the psychophysical case, in which there also seem to be two property changes: a change in the brain and a new feeling in the mind. The answer would be that whereas in the purely physical case there is an *imparting* of one thing to another, the brain state does not seem to impart anything to the mind. There is no transmission of any kind. The relationship between the mental and the physical state should rather be understood as simultaneous change concerning two necessarily conjoined properties.

## Causation and Activity/Passivity

The question of shared modes already introduced some of the problems around the relation of physical and mental states in psychophysical interaction. In this section I aim to go further in this direction and show the interpretative inadequacy of the standard understanding of this relation as straightforwardly causal, as well as of a clear division between the active and passive roles played by the two substances in psychophysical processes.

The question usually asked concerning the causal character of psychophysical processes in Descartes is whether the relationship between the mental and the physical state is causal or *short of* that. There is an occasionalist reading (Baker and Morris 1996): body and mind do not actually act as real efficient causes on each other (or at least the body does not really act on the mind); rather, the ultimate cause of change in the other substance is God. The textual evidence adduced will show, in contrast, that in some ways the relationship between mind and body in interaction is *more*, rather than less, than causal. Many things enter into causal relations with each other, and the fact that a particular mind and a particular body do so does not mark them out as belonging to each other in any special way. In the instance of a passion, however, the mental property is not simply caused by the bodily state, but also shows that body to be intimately connected to that mind. A mere causal characterization does not do justice to the intrinsic character of the mind-body relationship revealed in psychophysical states, the relationship that Descartes highlights with the disanalogy of the sailor and his ship in *Meditations*.

One of the ways in which Descartes goes about describing the relationship between mental and physical states involved in the passions may be called “semantic”. Descartes’ account of vision, a psychophysical operation, incorporates the idea of non-image-like *signification*. What Descartes actually describes is how an adequate amount of visual information can be retained through the process of vision without the assumption of “little images” travelling from the object to the eye. The psychophysical process is seen as transmission of information, with the help of signs that do not (fully) resemble what they signify. Things are “represented to the soul” by such non-resembling signifiers: “movements (both of the gland and of the spirits and the brain) *which represent certain objects to the soul*” (AT XI 369/CSM 1.348). The relationship of signification is not between the trace on the gland and the mental state, however, but between the trace on the gland and the thing it is the trace of *and* between that latter and the content of the mental state. The intentional object of the mental state is not a brain state. Moreover, the semantic approach tells us something about the *content* of the mental state in relation to the object of the passions, while it does not say much about the relationship between the mental and the physical properties from a *metaphysical* point of view.

The physical states immediately preceding or corresponding to mental states in the case of the passions are end states of long physical processes involving external organs, nerves, animal spirits and the brain. If we seek an account of the relation between the physical (end) state and the mental state, the whole of the process involving the interaction needs to be taken into consideration (although, of course, the end state is the most relevant physiological state from our point of view).

Let us begin with the account of sight in *Treatise on Man*. Light rays coming from points of external objects press upon the back of the eye. From there, optic nerves comprising tiny fibres stretch to the internal surface of the brain. When a point on the back of the eye is thus pressed, it pulls a whole fibre at the end of which the opening of a tube gets enlarged in the brain. In this way the figure traced on the back of the eye gets transmitted to the internal surface of the brain. As the tube enlarges, animal spirits from the pineal gland enter these in greater numbers. The leaving of the spirits from the surface of the gland traces the same figure as appeared on the back of the eye and the internal surface of the brain. Now, what is relevant for us: the soul “directly considers” the figure imprinted on the surface of the gland (AT XI 177/CSM 1.106). We don’t get an indication here of the way in which this last step, which seems to involve some activity of the soul, takes place.

In *Optics*, also written around 1630, Descartes makes another attempt at describing the functioning of the faculty of sight. He broaches the topic with criticism of the scholastic doctrine of images formed in external objects and transmitted to the sense organs, which appear as “little pictures [...] in our head” (AT VI 112/CSM 1.165). This teaching, while relying on the appeal of the idea that the subject has access to such pictures, fails to explain the process leading up to having them. The way to produce a convincing explanation opens up when the idea of thorough similarity (image-like-ness) all the way down is dispensed

with. We can be “stimulated by” things that do not resemble the items they signify.<sup>4</sup> The differences in the neural motions caused correspond to the differences in the qualities of the objects they are caused by; they do not have to resemble those qualities in an image-like way. Resemblance of images is replaced by imprinting as the focal idea of the discussion of sight.

Although “in order to have sensory perception the soul does not need to contemplate any images resembling the things which it perceives [...] the objects we look at imprint quite perfect images of themselves on the back of our eyes” (AT VI 114/CSM 1.166). There *is* an image imprinted that is similar to the object represented; the difference is that no item wanders from the object to the sense organs, but the similarity of the object and the imprinted image is mediated through motion that is dissimilar to both. The image formed in the back of the eye, just as in *Treatise on Man*, is further imprinted on the inside of the brain. As a last step, “the movements composing this picture act directly on the soul”, causing it to have sensations that correspond to these movements “as ordained by nature” (AT VI 130/CSM 1.167). Here the suggestion is more of passivity, rather than activity, of the mind. There is a movement in the brain that makes us aware of colour and light: “we must suppose our soul to be of such nature that what makes it have the sensation of light is the force of the movements taking place in the regions of the brain where the optic nerve-fibres originate, and what makes it have the sensation of colour is the manner of these movements” (ibid.). Other kinds of perception (hearing, tasting, etc.) are similar in this respect.

What can this “contemplation” or “consideration” be? We can hardly think that it is an intentional act directed at the relevant state of the brain: our transparent minds do not seem to be directed at *brain states* in sensation. Furthermore, the mind would not seem to be passive in such a state. Where does this fundamental uncertainty concerning the connection between the physical and the mental property in the passions come from?

In order to understand some of the reasons, we have to look back to the *Rules* and an inherited problematic Descartes was facing. In the *Rules*, “ideas” are of a corporeal nature. Cognitive functions are for the most part physical, thus can be, in some fashion, shared with animals. Corporeal functions in cognition are explained by analogies designed to make vivid the possibility of the transmission of information “pure and without body” (that is, without an entity actually passing from one place to the other) (AT X 414/CSM 1.41). The “pure intellect” has a rather restricted, somewhat superadded role (which can explain why *Treatise on Man* talks about the soul in a way that Descartes arguing against the sailor-in-the-ship view would not have).

Whether the sensory power resides in the mind or the body had been a controversial issue in the Middle Ages. In Descartes’ own time, the idea that sensation does not involve the intellect could be found, for instance, in Eustachius a Sancto Paulo (Eustache de Saint-Paul), whose *Summa Philosophiae* Descartes read and regarded highly.<sup>5</sup> In Eustachius, the

---

4 See a detailed treatment of this topic in Ben-Yami 2015.

5 See his letter to Mersenne, AT III 232.

sensory power operates through sense organs only.<sup>6</sup> When the intellect operates on material obtained by the senses, it “abstracts” the intelligible species from that material.

In contrast to Descartes’ earlier approach along somewhat similar lines, in his later philosophy the competences of the mind increase, and “ideas” become by definition mental. That is, sensory ideas, which used to be identified as brain images, are to be “mentalized”. At the same time, a very close link between sensory ideas and the brain states Descartes once identified with ideas obviously needs to be acknowledged. This relation sometimes seems to be closer to a “dual aspect” (sometimes also called “dual attribute”) conception of psychophysical occurrences.<sup>7</sup> As the passion of the mind is the action of the body and vice versa, as the mind “contemplates” or “considers” the body in a way that the mind itself remains passive (thus, it does not seem to produce a *separate* intentional act), and “movements” are said to “represent objects to the soul” (rather than cause a state in which objects are represented), we might want to try a dual aspect interpretation of psychophysical occurrences in Descartes. In this case, the relationship between the mental and the physical properties would be one of – qualified – identity rather than causation.

According to such an interpretation, the mind “contemplating” a change in the body would be that very change in the “idea mode”. The relation would not have to be one of representation, as the mind is not directed at brain states. Rather, in the case of external sensation, the mental aspect of the state would be the mind directed at the object that, through the organs of perception, caused the brain state. In the other, mind-to-body direction, such an interpretation can be made vivid by the example of attention, which is an action of the soul on the body. In the case of being attentive to something, our volition “keeps the gland leaning in one particular direction” (AT XI 361/CSM 1.344). It is not the case that there is an act of the will, followed by a bodily change; rather, there is one event, that of keeping the gland in a particular position, with two, mental and physical, “aspects”. If we want the mind to move the body, it causes a change in the pineal gland that is not different from the mental move of willing itself. On this reading, the one-to-one correspondence between mental and physical changes (the soul “receives as many different impressions – that is, it has as many different perceptions as there occur different movements of this gland” [AT XI 354/CSM 1.341]) and is like that between two sides of a coin, rather than a coordination between two independently individuated events.

This reading gives us a better grasp of the pineal gland as genuinely connecting brain processes with mental contents. For it would mean that the one movement of the gland is both directly physical and mental. If, in contrast, on a two-event interpretation, there is a physical event involving the pineal gland followed by a mental event involving the pineal gland, there would seem to be two different changes to the gland itself, with the connection between them being rather unclear. We might even think that it requires a third,

---

6 See Schmalz 1997, 47–8.

7 Originating from the interpretation of Spinoza, “dual attribute” is meant to convey the idea that there are not two kinds of substance, mind and body, but a single one, with mental and physical aspects. The use of the phrase is not so constrained, however: it is also used with *events* or *states* with dual properties.

“conversion” event that does the alleged mediation, for neither the physical nor the mental event is one of mediation or transmission of information.

It could be objected that the pineal gland is, after all, a physical organ. The movements of a physical organ are physical, and the mind only feels what happens in that organ, but does not have anything to do with its movements; there is no “mental” side to those movements. But if this is so, how does the soul “directly exercise its functions” through that organ? How is the pineal gland “the principal seat of the soul” (AT XI 352/CSM 1.340)?

Whether the semantic conception is reconcilable with the causal view, they are not the same in Descartes. Nor is the “natural institution view” quite the same as either of these. Descartes seems to be trying, unsuccessfully, to offer an account of the relationship between mental and physical states that goes beyond, but also accommodates, causation. As I have tried to show, his account of the relationship between mental and physical states sometimes shows more affinity to a dual aspect theory than to straightforward dualism.

Here, we apparently reach the limits of interpretation. It does not seem like there is a way to make a decisive choice between possible interpretations, for Descartes’ position is, I believe, inherently ambiguous. He is trying to make the causal character of the mind-body relationship convincing; at the same time, he is also trying to make out that relationship to be something *more*. Critics have been largely concerned with and attacked the former; the latter is perhaps too little understood to properly address and criticize.

We can safely conclude, however, that, in the case of the passions, the picture is not simply that of (a change in) the active body causing a subsequent change in the passive mind, for (1) activity seems more evenly distributed, and (2) the view that there is an event in the body which is followed by an event in the mind just does not seem reconcilable with the one voiced by Descartes that action and passion are a “single thing” and (3) the distinction between the bodily and the mental event or property is much less straightforward.

## Bibliography

### Primary Sources

- AT Adam, Charles and Tannery, Paul (eds). 1996. *Oeuvres de Descartes*, 11 Vols. Paris: Vrin.
- CSM Cottingham, John, Stoothoff, Robert and Murdoch, Douglas (trans). 1985. *The Philosophical Writings of Descartes*. Vols. 1–2. Cambridge: Cambridge University Press.

### Secondary Sources

- Ben-Yami, Hanoch. 2015. *Descartes’ Philosophical Revolution: A Re-Assessment*. London: Palgrave
- Baker, Gordon and Morris, Kathrine. 1996. *Descartes’ Dualism*. Oxford: Oxford University Press.

- Clatterbaugh, Kenneth. 2009. "The Early Moderns." In *The Oxford Handbook of Causation*, edited by H. Beebe, K. Hitchcock and P. Menzies, 55–72. Oxford: Oxford University Press.
- Schmaltz, Tad M. 1997. "Descartes on Innate Ideas, Sensation, and Scholasticism: The Response to Regius." In: *Studies in Seventeenth-Century European Philosophy*, Oxford Studies in the History of Philosophy, vol. 2, edited by M. A. Stewart, 33–73. Oxford: Oxford University Press.
- Schmaltz, Tad M. 2008. *Descartes on Causation*. Oxford: Oxford University Press.
- Wilson, Margaret D. 1978. *Descartes*. New York: Routledge.

## Leibniz on Personal Identity

Leibniz's view on personal identity has been the object of numerous discussions and various interpretations. Among others, the controversies revolve around the following questions: (1) What is the relation of Leibniz's conception to the Cartesian view on personal identity? Is it a completely new idea or some modification of Descartes'? (2) To what extent do Locke's ideas, especially Locke's distinction between being the same substance, organism and person, lay the basis for Leibniz's conception of personal identity? (3) What role does psychological continuity play in Leibniz's conception of personal identity? Did he indeed claim that a person's identity cannot solely arise out of sameness of substance? (4) Is Leibniz's solution to the problem of personal identity compatible with his deepest metaphysical commitments? Can it be seen as a conclusive solution to the problem? (5) Is Leibniz's effort to offer an account of personal identity by combining the substance-oriented view with the psychological view a coherent solution? (see: Mijuskovic 1975; Scheffler 1976; Curley 1982; Wilson 1999; Jolley 1984; Vailati 1985; Mates 1986; Thiel 1998; Noonan 1989; Bobro 2004).

Before I specify which of the above problems come into focus in my work, let me refer to three opinions formulated by Samuel Scheffler, Margaret Wilson and Ezio Vailati. A presentation of their views will allow me, first of all, to highlight why Leibniz's view on this issue leads to so many controversies, and secondly, to indicate the points where my interpretation departs from those of other researchers, especially from the ones offered by M. Wilson and S. Scheffler.

# 1.

In Samuel Scheffler's opinion – whose text opened the debate on the subject – Leibniz did not manage to demonstrate why memory and other psychological phenomena are insufficient to establish the identity of persons over time. Scheffler suggests that Leibniz limits himself solely to the statement that only the so-called *a priori* reasons which result from the continued existence of the same substance are a sufficient basis for being *the same* person over time (Scheffler 1976, 223). According to Scheffler, Leibniz's writings fail to provide any substantial arguments for accepting this claim. Scheffler claims Leibniz's only argument consists of a fairly vague conviction that accepting memory or other psychological phenomena as a condition of personal identity is at variance with our natural intuitions. If one agreed that personal identity is based on the continuity of memory, one would have to acknowledge that the complete loss of memory (e.g. as a result of an unfortunate accident) would result in the loss of personal identity. And this is exactly what – according to Leibniz – is supposed to be at odds with our natural intuitions. According to Scheffler, this argument is not convincing since the proponent of memory as a necessary condition of personal identity, while investigating cases of memory loss or gaps, could claim that in such cases our natural intuitions do not yield credible judgements when it comes to personal identity. It is not memory as a condition of identity that must be rejected but our common sense convictions in this respect. Leibniz failed to provide any argument against this line of thought.<sup>1</sup> As Scheffler suggests, Leibniz seems not to have had a clear picture of his own understanding of personal identity (Scheffler 1976, 239).

Equally severe criticism of Leibniz's position was levelled by Margaret Wilson. As she states, Leibniz failed to formulate a coherent and uniform theory of personal identity. His stance – interesting and important as it might be – contains many inconsistencies, therefore it cannot be adopted as a satisfactory solution to the problem. Wilson claims that, especially in *New Essays Concerning Human Understanding*, it is difficult to state unambiguously what Leibniz believed to be the basis of being the same person over time. After

---

1 “Of course, Leibniz is aware of the importance of the psychological phenomena, and he even says that they are what convince us *a posteriori* that certain time-life slices are part of our lives, and that others are not. Yet in seeking a suitable reason in virtue of which it is the case that a person retains identity over time, Leibniz does not even consider memory and related phenomena as candidates. He simply asserts that ‘there must [...] be some reason *a priori*’, and proceeds to locate the reason in his theory of substance. In arguing from the premise that there must exist a reason why I am the same person this week as last to the conclusion that the reason must derive from the concept of substance and be knowable *a priori*, Leibniz reveals that he simply refuses to count phenomena like memory and continuity of consciousness as reasons of the relevant kind. Yet in the abruptness of his transition from premise to conclusion, Leibniz further reveals that he has brought no arguments to bear in support of his view that memory and related phenomena are insufficient reasons. Relying on an intuition that mere psychological continuities couldn't be enough to bind sequences of time-life slices into lives, Leibniz simply begs the question by assuming the correctness of his intuition and failing even to produce arguments against the sufficiency of psychological criteria.” (Scheffler 1976, 224–5)



Descartes, he holds that the continued existence of substance – i.e. the existence of a soul or “I” – is what ultimately determines identity. Simultaneously, contrary to Descartes, he emphasizes that the preservation of psychological continuity based on self-consciousness and memory seems to be indispensable to the moral and religious significance of personal identity. This, in turn, reduces the distance between Leibniz and Locke. The latter believed that psychological continuity (i.e. memory and consciousness) is crucial for being the same person. It is because of an attempt to combine these two positions that Leibniz runs up against difficulties.

According to Wilson, the most serious difficulty that the reader encounters in Leibniz’s texts, especially in *New Essays Concerning Human Understanding*, lies in that, while identifying the notion of a person with spiritual substance, Leibniz does not exclude, at least logically, the possibility of altering spiritual substance while preserving personal identity based on the continuity of one’s psychological life. As a result, the continuity of spiritual substance’s existence turns out to be an unnecessary basis of personal identity, even though it is the ontic core of a being, which seems to be a glaring inconsistency (Wilson 1999, 380; Jolley 1984, 135–8). Moreover, as Wilson continues to explain, the knowledge about substance, which we gain through our internal experience of ourselves, is characterized by Leibniz differently in various places. Some of his texts suggest that this knowledge gives incontrovertible proof of the substantiality of one’s “I”, but also of one’s authentic individuality (Gr, vol. 2: 557–58). Others, in turn, promote the view that the knowledge we gain through our internal experience, important as it might be, does not constitute the whole content contained in the idea of individual substance, identical with an individual concept (L 332–33). As a result, it cannot be regarded as a sufficient ground for personal identity (Wilson 1999, 381).

A different point of view in the discussion was outlined by Ezio Vailati. As Vailati explains, such severe criticism of the results of Leibniz’s studies of personal identity is ill-founded. Contrary to Margaret Wilson and Samuel Scheffler, he believes that Leibniz’s statements in this respect are not ambiguous at all and that the theory of personal identity which emerges from them is not incoherent. Vailati demonstrates this with the following three points. In the first place, it is not true that Leibniz defines a person in *New Essays Concerning Human Understanding* only on the basis of their substantial principle, i.e. soul and mind. A person for Leibniz is also a moral and religious being. This is where the requirement to combine substantial continuity and psychological continuity (consciousness and memory) as the condition of personal identity comes from. Secondly, Leibniz clearly explains that potential separation of somebody’s consciousness and memory from their substantial principle is possible only logically. It is at odds, though, with “the order of nature”. From the point of view of nature, such a situation is ruled out (Vailati 1985, 38). Thirdly, it is necessary to bear in mind that for Leibniz consciousness establishes personal identity insofar as it is “accompanied by truth”, that is when consciousness is veridical.<sup>2</sup>

---

2 “As regards self, it will be as well to distinguish it from the appearance of self and from consciousness. The self makes real physical identity, and the appearance of self, when accompanied by truth, adds to it personal identity.” (NE: 2, 27, § 9).

When these three points are taken into account, the lack of clarity of Leibniz's conception of personal identity disappears.<sup>3</sup>

## 2.

Broadly speaking, the previous statements are the presentation of views on Leibniz's conception of personal identity. Clearly, the first two are very critical. Both Scheffler and Wilson believe that they managed to reveal some essential difficulties and mistakes present in Leibniz's theory, thus disqualifying it. The real question is whether the above criticism pertains to what can be found in Leibniz's texts. It seems to me that Leibniz's theory combining the substantial basis of personal identity with the demand of psychological continuity is not burdened with any particular inconsistency. Considering what Vailati presented, there is a possibility of interpreting Leibniz's statements which dismisses Margaret Wilson's accusation of Leibniz's reported inconsistency. What needs to be observed is that Leibniz does not limit the notion of personal identity to a substantial principle, i.e. the soul or "I", which comes to pass in Descartes' doctrine.

If Vailati is right (and there are a number of reasons to believe so) there emerges a basis on which the consistency of Leibniz's position can be defended. This leads to another question: can Leibniz offer arguments powerful enough to support his two theses whose truth he was trying to prove? What I mean here are arguments which support the claim that, on the one hand, personal identity should be treated as a structure which consists of two layers, i.e. substantial and psychological, and, on the other, that the ultimate basis of personal identity consists in the continuity of substance, even if substantiality fails to exhaust the concept of "personal identity". In my opinion, Leibniz did present such arguments. That is why I believe that it is inadequate on Scheffler's part to claim that, apart from a vaguely characterised intuition, Leibniz does not advance arguments which support his hypothesis that substantial continuity is of fundamental importance for the preservation of personal identity. It is my conviction that a more in-depth analysis of Leibniz's texts leads to the conclusion that the so-called intuition is not his only argument which proves that the continuity of psychological phenomena alone does not guarantee the preservation of personal identity.<sup>4</sup>

I also think that Wilson's view that if Leibniz uses both Descartes' and Locke's ideas in *New Essays Concerning Human Understanding*, his position on personal identity is only

---

3 "The charge of inconsistency arises from attributing two incompatible theses to Leibniz, namely, that I could continue as a person without my mind and that I, as a person, am my mind. There is no doubt that Leibniz accepts the former claim, albeit with the non-trivial proviso that 'could' must express merely a logical and not also a natural possibility. But it should also be clear from what I have argued above that for Leibniz I, as a person, am not my mind or my substantial self, although in the order of nature I cannot be separated from it." (Vailati 1985, 43)

4 A similar opinion is expressed by other authors: Noonan 1989, 46; Jolley 1984, 143–4.

a compilation of what these two authors assert is groundless. It fails to recognize the fact that, already between 1680 and 1690, i.e. long before he became acquainted with Locke's theory, Leibniz pointed out that personal identity should be approached from two points of view: metaphysical and psychological.<sup>5</sup> Most importantly, Wilson's opinion ignores the fact that Leibniz differed from Locke in his understanding of the continuity of consciousness of past experiences as a condition of personal identity. First of all, Leibniz was not as rigid as Locke in his view on the continuity between consciousness and memory. Referring to cases of memory gaps when a person loses consciousness or goes into a deep sleep, Leibniz observed that such cases do not destroy somebody's identity. Because of this, as Leibniz believed, it is sufficient if there is a connection based on consciousness between two neighbouring states – even if there is a gap between them caused by memory loss – in order to preserve psychological continuity (NE: 2, 27, § 9). Additionally, and contrary to Locke, Leibniz held that one can refer to accounts of other people to preserve personal identity. Here, he pointed to cases of long-term amnesia, when memory gaps are filled with false or accidental content, which are accompanied, however, with the conviction of accuracy (confabulation) and cases of distorted memory (paramnesia). The fact that such situations are actually the case must make one believe that not only direct consciousness but also other people's accounts can be of importance for the preservation of personal identity.<sup>6</sup>

It is worth emphasising here that bearing in mind accounts of third parties assumes that, according to Leibniz, the body can also have some importance for the preservation of personal identity. As Leibniz explained, even though the human body is not the essence of a person, it is one way of fulfilling the relation of one human being to other beings in the world. It is through bodies that the coexistence (mutual subordination) of all individual beings in the world is possible.<sup>7</sup> What is more, the beginnings of a body are, so to say, predetermined and permanently connected with particular human beings.<sup>8</sup> That is why

---

5 See: *Discourse on Metaphysics* and *Correspondence with Arnauld*.

6 “Thus, if an illness had interrupted the continuity of my bond of consciousness, so that I didn't know how I had arrived at my present state even though I could remember things further back, the testimony of others could fill in the gap in my recollection. I could even be punished on this testimony if I had done some deliberate wrong during an interval which this illness had made me forget a short time later. And if I forgot my whole past, and needed to have myself taught all over again, even my name and how to read and write, I could still learn from others about my life during my preceding state; and I would have retained my rights without having to be divided into two persons and made to inherit from myself! All this is enough to preserve the moral identity that makes the same person.” (NE: 2, 27, § 9).

7 “For although monads are not extended, they nevertheless have a certain kind of situation [*situs*] in extension, that is, they have a certain ordered relation of coexistence with others, namely, through the machine which they control. I do not think that any finite substances exist apart from a body and that they therefore lack a position or an order in relation to the other things coexisting in the universe. Extended things involve a plurality of things endowed with position, but things which are simple, though they do not have extension, must yet have a position in extension, though it is impossible to designate these positions precisely as in the case of incomplete phenomena.” (*Correspondence with de Volder* [20 June 1703], in L, 532).

8 Erdmann, 653–63.

neither complete birth nor complete bodily death exist and what we describe as “*generation* is a development and an increase, just as what we call death is an envelopment and a diminution”.<sup>9</sup> Furthermore, the soul expresses its own body directly and more visibly than other bodies. It expresses other bodies indirectly – through its own body.<sup>10</sup> All this, as Leibniz believed, enables us to take into account the role of the body in cases we discussed above

Taking the above considerations into account, I would like to address two issues in the remaining part of the essay. First, I will discuss the reasons which Leibniz listed to support his thesis that personal identity requires both the continuity of substance and the continuity of some psychological phenomena. Then, I will turn to identifying Leibniz’s arguments which support the thesis that what ultimately provides a person with identity is their substantial principle, i.e. the soul or “I”.

### 3.

Leibniz presented his position on personal identity mainly in *Discourse on Metaphysics* (1686), *Correspondence with Arnauld*, (1686–87), and, most importantly, in *New Essays Concerning Human Understanding* (Book 2, chapter 27). Important comments related to the notion of personal identity are also included in *Monadology* (1714) and *The Principles of Nature and of Grace, based on Reason* (1714). In both texts, Leibniz points out that the theory of personal identity must fulfil two fundamental tasks. Firstly, it must identify a factor which guarantees permanence, coherence, internal cohesion and order of individual changes which a person experiences over time. Secondly, the theory of personal

---

9 *The Monadology*, § 73 in L, 650. Here, Leibniz supported his claim with the theory called “perforation”, which dominated in the science of the seventeenth and eighteenth centuries. According to this theory, the embryonic development consists in the growth of a fully-developed, miniature being, which is located either in an egg or in a spermatozoon. It was replaced with the theory of epigenesis – documented by Ch. F. Wolff in 1759 – which is currently adopted in embryology. As the theory goes, the development consists in gradual differentiation of cells that are created after the zygote is divided. The next stages involve the creation of tissues, organs and systems.

10 “Thus, although each created monad represents the whole universe, it represents more distinctly the body which is particularly affected by it and of which it is the entelechy. And as this body expresses the whole universe by the connection between all matter in the plenum, the soul also represents the whole universe in representing the body which belongs to it in a particular way.” (*The Monadology*, § 62 in L, 649). This view presented by Leibniz was criticized by Arnauld. According to Arnauld, if our soul expressed its own body directly and more clearly than other bodies, it should be aware of numerous bodily processes such as digesting, nourishing. Yet it does not have this knowledge (See: *Arnauld’s letter to Leibniz*, dated 4 March 1687). In response to this accusation, Leibniz asserted that his position does not require the awareness of all bodily processes. The point, Leibniz continued to explain, is that changes happening in our body are perceived faster by our soul than by external changes. (See: *Leibniz’s letter to Arnauld*, dated 9 October 1687).

identity must identify a factor which can lay down the principles behind both the moral understanding of a human being and the religious sense of their immortality (L, 325). Only in this way does the theory of personal identity stand a chance of providing a correct and precise answer to the question ‘what makes a given person the same person regardless of changes over time’. According to Leibniz, the first task can only be fulfilled by referring to the continuity of substance (the soul or “I”). In order to fulfil the second task, one must resort to psychological continuity.<sup>11</sup>

Leibniz’s theory of a person is founded on the conviction that a person is an entity composed of two aspects: the metaphysical one, rooted in the world of nature, and the moral and religious one, rooted in the world of grace (*Discourse on Metaphysics*, § 34, 35 in L, 325–26). In the case of a person, these two aspects are mutually adjusted, even though they cannot be reduced to one another. That is why each of them requires the application of different principles that guarantee being the same person.<sup>12</sup>

In this context, it is clear why Leibniz distances himself both from Descartes’ and Locke’s solutions to the problem of personal identity. It can be argued that his main reservation against Descartes’ and Locke’s conceptions did not concern what they claim, but what they omit. As far as Descartes is concerned, Leibniz agreed undoubtedly that personal identity relies on the continuity of somebody’s substance, i.e. the soul (“I”). However, he believed that, focusing on substantiality, Descartes disregarded the impact of psychological continuity (*Discourse on Metaphysics*, § 36 in L, 326–27). In the meantime, without consciousness and the memory of what a person was, the sensitivity to punishment and reward – which is a necessary condition of the existence of moral qualities – is impossible. What is more, the conception according to which a person is constituted only by their substance is at odds with the doctrine of immortality, since it strips the idea of immortality of content which is important for the idea from the point of view of ethics and religion. It thwarts the legitimacy of any compensation, any punishment and any progress towards higher excellence. It even seems that immortality without the consciousness of past experiences would not make any sense.<sup>13</sup> Immortality is not the same as the continuous existence

---

11 “I also hold this opinion that consciousness or the sense of I proves moral or personal identity. And that is how I distinguish the unendingness of a beast’s soul from the immortality of the soul of a man: both of them preserve real, physical identity; but it is consonant with the rules of God’s providence that in man’s case the soul should also retain a moral identity that is apparent to us ourselves, so as to constitute the same person, which is therefore sensitive to punishments and rewards.” (NE, 2, 27, § 9).

12 “But in order to support by natural reasons the view that God will preserve for all time not merely our substance but also our person, that is to say, the memory and knowledge of what we are (though the distinct knowledge is sometimes suspended in sleep and in fainting fits), we must add morals to metaphysics.” (*Discourse on Metaphysics*, § 35 in L, 326).

13 “I therefore assert that the immortality of soul, as established by Descartes, is useless and could not console us in any way. For let us suppose that soul *is a* substance and that no substance perishes; given that the soul would not perish and, in fact, nothing would perish in nature. But just as matter, the soul will change in its way, and just as the matter that composes a man has at other times composed other plants and animals, similarly, this soul might be immortal in fact, but it might pass through

of the same soul, but it assumes the continuity of the same personality.<sup>14</sup> For these reasons, Leibniz maintained, psychological continuity is a necessary condition of the preservation of personal identity.

As for Locke, Leibniz agreed with him on the point that if a person did not preserve the consciousness of past experiences, they would not be able to be the same person from the moral and religious point of view.<sup>15</sup> However, it would be not the proper conclusion, as Leibniz claimed, that personal identity can be preserved without the reference to substantial continuity. According to Leibniz, for numerous reasons (which I present below), one should persist in thinking that only due to one substance, various manifestations of a person over time can constitute one, authentic whole, and it is neither the continuity of self-consciousness nor the continuity of the memory of past experiences (Thiel 1998, 899).

Leibniz concluded that, instead of looking for one foundation of personal identity, it is significantly more reasonable to assume that, in the case of a person, two dimensions of identity are equally important: the continuity of substance (the soul or “I”) and the continuity of consciousness and memory.<sup>16</sup> The continuity of substance is the so-called real or metaphysical identity, while the continuity of consciousness and memory is the so-called moral identity or identity “that is apparent to the person”. Real identity and moral identity

---

a thousand changes without remembering what it once was. But this immortality without memory is completely useless to morality, for it upsets all reward and punishment.” (*Letter to Molanus* in AG, 243).

- 14 “Hence, though animals may pass through a thousand transformations like that which we see when a caterpillar changes into a butterfly, yet from the moral or practical point of view the result is just as if they had perished; indeed, one may even say that they have perished in a physical sense, that is, in the sense in which we say that bodies perish through their corruption. But the intelligent soul, knowing what it is and being able to say this little word ‘I’ which means so much, not merely remains and subsists metaphysically (which it does in a fuller sense than the others) but also remains the same morally and constitutes the same character. For it is memory or the knowledge of this ‘I’ which makes it capable of punishment and reward. Likewise, the immortality which is demanded in morals and religion does not consist merely in this perpetual subsistence which is common to all substances, for without a memory of what one has been, there would be nothing desirable about it.” (*Discourse on Metaphysics*, § 34 in L, 235). “But the fact is that they confused indestructibility with immortality, whereby is understood in the case of man that not only the soul but also the personality subsists. In saying that the soul of man is immortal one implies the subsistence of what makes the identity of the person, something which retains its moral qualities, conserving the *consciousness*, or the reflective inward feeling, of what it is: thus it is rendered susceptible to chastisement or reward. But this conservation of personality does not occur in the souls of beasts: that is why I prefer to say that they are imperishable rather than to call them immortal.” (*Theodicy*, § 89: 175).
- 15 According to Noonan, the fundamental affinity between Leibniz and Locke lies in the fact that both Leibniz and Locke regarded “the person” as a “forensic term” and both were “vividly aware of the need to give an account of personal identity which makes comprehensible why it matters.” (Noonan 1989, 46). See also: Jolley 1984, 141; See: *Discourse on Metaphysics*, § 35 in L, 326.
- 16 “As regards [to the] self, it will be as well to distinguish it from the appearance of self and from consciousness. The self makes real physical identity, and the appearance of self, when accompanied by truth, adds to it personal identity. So, not wishing to say that personal identity extends no further than memory, I want even less to say that the self, or physical identity, depends on it.” (NE, 2, 27, § 9: 111).



– as Leibniz further explains – are, in a logical sense, two separate structures of personal identity. Each of them has different foundations. The foundation of the former is the continuity of individual substance, while the latter is rooted in the continuity of consciousness and memory.<sup>17</sup>

According to Leibniz, this does not mean that the continuity of the same substance is unimportant for the moral identity of a person, since it originates in the continuity of consciousness and memory, that is in the source which is independent of the substance. As Leibniz explains, a logical possibility (i.e. something that is not internally contradictory) must be differentiated from a natural possibility, i.e. something that can be accepted by the order of nature. While changing the spiritual substance without modifying personal identity based on the psychological life of a person is feasible according to a logical possibility or from the point of view of the absolute power of God, it is unacceptable under a natural possibility. From the point of view of the order of nature, it is impossible for the continuity of consciousness and memory to be preserved, no matter if it belongs only to one substance or to a number of consecutive substances, as is the case in Locke's doctrine.<sup>18</sup>

In Leibniz's view, arguments that prove the existence of such interdependence are provided already by our internal experience, in the light of which we learn about a close and strong connection between our psychological life and "I" as individual substance. Even if this experience, as Wilson observes, is not a decisive argument in Leibniz's system, still he regards it as a vital piece of data.<sup>19</sup> Whatever it makes us realize undermines the view

---

17 This shows that moral identity cannot be seen merely as an epistemic condition of real identity, which is based on identity of substance. Although some of Leibniz's formulations may suggest such interpretation, considering his whole discussion of personal identity, it becomes evident that he rather saw consciousness and memory as constitutive of identity. See Thiel, *Personal Identity*, 900. See also *Discourse on Metaphysics*, § 34 in L, 325: "But the intelligent soul, knowing what it is and being able to say this little word 'I' which means so much, not merely remains and subsists metaphysically (which it does in a fuller sense than the others) but also remains the same morally and constitutes the same character. For it is memory or the knowledge of this 'I' which makes it capable of punishment and reward."

18 According to U. Thiel, in this lies, "the most fundamental difference between Leibniz and Locke: for Locke it is a real possibility that there be personal identity without substantial identity [...]. To Leibniz, however, this is a mere logical possibility. It 'would be a miracle': it would 'disrupt the order of things for no reason, and would divorce what can become before our awareness from the truth – the truth which is preserved by insensible perceptions' (NE, 27, § 18). According to the 'order of things', Leibniz argues, real identity must be presupposed by apparent identity. Thus, although he does not equate personal with substantial identity, he holds that the former depends on the latter. Whereas Locke argued for keeping personal and substantial identity separate, Leibniz maintained what was assumed by the Cartesians, namely, that the (personal) identity required for morality can be preserved only by the metaphysical identity of the self as immaterial soul." (Thiel 1998, 902).

19 "That we are not substances is at variance with experience since we actually gain the knowledge about substances only on the basis of the most internal experience of ourselves, when we get to know our own 'I' and – using this equation – we ascribe the name of a substance to God and other monads." (Gr, 2: 557–8). See also: GP, 6: 499–508. This argument plays a prominent role in the doctrine. One's own "I" is the paradigm of the general idea of a substance. A soul (mind) is precisely the point – according to many historians – which contains the source of the idea of a substance (monad) as a non-spatial, complete and substantial being; the idea which is the generalization of criticism levelled at Spinoza's

that the consciousness of past experiences alone, irrespective of a substantial principle, guarantees being the same person sufficiently. What is more, according to Leibniz, personal identity can be based on consciousness alone only if a man could be a mere machine and still possess consciousness (NE, 2, 27, § 9). In the meantime, the hypothesis about a machine possessing consciousness is, in fact, at variance with the natural order of things. The self-consciousness of a person embraces both the direct self-consciousness of individual experiences of a person and the direct self-consciousness of one's own "I" as a subject. A person knows oneself (one's "I") directly. The consciousness of "I" accompanies all perceptible experiences of a person who goes through them. In each of these experiences "I" is given as a whole without any conception of its parts. Direct consciousness shows "I" as a singular, simple, not complex, immaterial being. A machine is essentially an aggregate, something complex, bodily. As a result, it is not possible for an aggregate to perceive itself as a singular "I" (Vailati 1985, 38). In Leibniz's view, the reflection itself is possible only if the subject of this activity is an entity capable of being over time. "I" perceives its thought *p* only after it happens, i.e. thought *p* is previous in terms of time to the consciousness of it, even though the interval between a thought and its consciousness can be so short that we are not aware of it. This fact does not pose any problems if "I" is a substance, i.e. *unum per se*, since a substance is the kind of being which preserves its identity over time. If "I" is not a substance but a machine which cannot be identical outside some moment, then "I" – as a machine – cannot practically refer to its thoughts in its reflections, since its identity does not go beyond the identity it has in this particular moment.

#### 4.

Let us now give a more detailed description of the reasons why Leibniz thought that the identity of substance must be accorded primacy in constituting personal identity, despite the considerable significance of consciousness for being the same person.

In *Discourse on Metaphysics*, in *Correspondence with Arnould* and in *New Essays Concerning Human Understanding*, Leibniz maintains that the concept of identity essentially refers to something that is simple. As for what is complex, identity is apparent or, at most, is

---

monism. The psychological origin of this thought is visible in Leibniz's writings clearly enough, but its metaphysical dimension is equally visible. On numerous occasions, Leibniz writes that mind is not only a psychological but also a metaphysical concept. The discovery of "I" is the idea we need to ponder in order to finally reach the real objective sphere, thoughts about other ideas, whose discovery cannot be guaranteed by external (sensual) cognition or operational skills. Leibniz comments: "It is also by the knowledge of necessary truths and by their abstractions that we rise to *reflective acts*, which enable us to think of what is called *I* and to consider this or that to be in us; it is thus, as we think of ourselves, that we think of being, of substance, of the simple and the compound, of the immaterial, and of God himself, conceiving of that which is limited in us as being without limits in him. These reflective acts provide us with the principal objects of our reasonings." (*Monadology*, § 30 in L, 646).



a matter of degree.<sup>20</sup> That is why, if we do not refer to a spiritually simple substance, the same person will never exist in the strict meaning of the word.<sup>21</sup> What is more, only by accepting the assumption of substantial continuity – the so-called *a priori* reason – an internal connection between different experiences of a person can be explained. This is how Leibniz elaborates on this thought in his letter to Arnauld:

Let there be a straight line ABC representing a certain time. And let there be an individual substance, for example, I, enduring or subsisting during that time. Let us first take me subsisting during time AB, and then me subsisting during time BC. Then, since the assumption is that it is the same individual substance that endures throughout, or rather that it is I who subsists in time AB, being then in Paris, and that it is still I who subsists in time BC, being then in Germany, there must necessarily be a reason allowing us truly to say that we endure, that is to say that I, who was in Paris, am now in Germany. For if there were no such reason, we would have as much right to say that it is someone else. It is true that my internal experience convinces me *a posteriori* of this identity; but there must also be an *a priori* reason. Now, it is not possible to find any reason but the fact that both my attributes in the preceding time and state and my attributes in the succeeding time and state are predicates of the same subject – they are in the same subject. (*Remarks on Arnauld's Letter* in AG, 73)

As can be seen in the excerpt above, the preservation of the unity and permanence of a person over time depends on the sameness of one subject (substance). A person preserves one's unity and permanence so long as all of one's properties are the properties of the same subject. It means that a person O2 existing at time t2 is the same person as a person O1 existing at time t1 as long as the properties of a person O2 existing at time t2 are the properties of the same subject as the properties of a person O1 existing at time t1. That is why, if one did not assume a permanent substantial subject, in which all properties are rooted, there would be no reason either for their interconnections or the authentic principle of unity between them. The unity and permanence of a person – from the beginning to the end of one's existence – can constitute themselves only thanks to the fact that consecutive properties belong to the same substantial subject.

---

20 "So we must acknowledge that organic bodies as well as inorganic ones remain 'the same' only in appearance, and not strictly speaking. It is rather like a river whose water is continually changing, or like Theseus's ship that the Athenians were constantly repairing. But as for substances that possess in themselves a genuine, real substantial unity, substances that are capable of actions that can properly be called 'vital', substantial beings [...] that are animated by a certain indivisible spirit, one can rightly say that they remain perfectly the same individual in virtue of this soul or spirit that constitutes the I in substances that think." (NE, 2, 27, § 4: 108–9).

21 "If plants and brutes have no souls, then their identity is only apparent, but if they do have souls their identity is strictly genuine, although their organic bodies don't retain such an identity." (NE, 2, 27, § 5: 109).

Apart from this, Leibniz proves that all experiences of a person – both perceptible for consciousness (sensible perceptions, awarenesses) and non-perceptible for consciousness (insensible perceptions) – are included in the individual substance of each person, which is one’s permanent subject: “I”. While a person can lose consciousness, i.e. apperception of some of one’s experiences, one cannot be entirely stripped of one’s perceptions. In the end, it is the continuity of perceptions and interconnections between them that decide about the sameness of a person.

An immaterial being or spirit can’t be stripped of all perception of its past existence. It retains impressions of everything that has previously happened to it, and it even has presentiments of everything that will happen to it; but these states of mind are mostly too tiny to be distinguishable and for one to be aware of them, although they may perhaps grow some day. It is this continuity and interconnection of perceptions that make someone really the same individual; but our awarenesses – i.e. when we are aware of past states of mind – prove a moral identity as well, and make the real identity appear. (NE, 2, 27, § 14)

Advancing the thesis about the necessity to accept the substantial foundation of personal identity, Leibniz did not focus only on metaphysical considerations. He believed that there were also other reasons for thinking about personal identity from the perspective of substantial continuity.<sup>22</sup> As he claimed, the assumption that it is possible to stop being the same person only by virtue of not having a direct consciousness of one’s experiences leads to absurd consequences and it is at odds with the natural conviction for several reasons. Firstly, if personal identity was based only on consciousness, its complete loss by a given person (e.g. as a result of an unfortunate accident) would mean the loss of personal identity. A person before and after an accident would be a completely different person. Secondly, if consciousness and memory were the only way personal identity can constitute itself, they would actually be all that constitutes a given person. It would lead to absurdity in the case when memory gaps were filled with false content (NE, 2, 27, § 9).

Then, he criticized the theory of personal identity without a substantial foundation of internal contradictoriness. Let us assume, Leibniz elaborated in *New Essays Concerning Human Understanding*, that in some other part of the universe there is a globe which does not differ sensibly from the earthly globe, which is inhabited by us, and that every person who lives there does not differ sensibly from any person on Earth who corresponds to him. This is how there are around four billion pairs of people with the same experiences and consciousness. Is each of these pairs one person or two people? It is not clear how to prevent the absurdity of the claim that two persons who live on two similar but infinitely remote globes are one and the same person based on the hypothesis that personal identity shall be decided solely by memory and consciousness without the need to refer to identity, or the diversity of the substance, or even without what appears to others (NE, 2, 27, § 23).

---

22 Leibniz formulates these arguments mainly in his discussion with Locke.

Apart from this, Leibniz accused the above theory of misinterpreting certain practical situations. According to Locke, the theory of personal identity as the continuity of self-consciousness and consciousness of a person's past experiences is universally in agreement with the practice adopted by legislators and judges. This is exactly why human law does not punish a madman – as Locke believes – for the deeds of a sane person and *vice versa*, since it treats them as two separate persons. For Leibniz, this conception was totally erroneous. The essence of law is to threaten to punish any wrongdoing in order to prevent it. But an insane person is unable to recognize the significance of the rigour of punishment, which is why in that situation the law refrains from punishing a person for what he or she did when being sane. What legislators do in such circumstances, then, does not result from the fact that a given human being is regarded as two persons, but from the fact that the same person is now unable to accept (understand) the rigour of punishment (NE, 2, 27, § 20).

According to Leibniz, all of these arguments establish a sufficient basis to recognize the continuity of substance as a necessary condition of personal identity. Only the claim that the substance of a person is the condition of personal identity saves the conception of a person from the aforementioned problems. Obviously, it must not be concluded that a person is entirely limited to their substance. The continuity of substance is only the necessary condition pertaining to a personal being. One additional condition of personal identity is moral continuity, which, in contrast to real identity, is based on the continuity of self-consciousness (reflection) and consciousness of past experiences (memory). Here is how Leibniz himself summed up his analysis in *New Essays Concerning Human Understanding*:

I have shown you the basis of true physical identity, and have shown that it doesn't [sic] clash with moral identity or with memory either. And I have also shown that although moral identity and memory cannot always indicate a person's physical identity, to the person in question or to his acquaintances, they never run counter to physical identity and are never totally divorced from it. Finally, I have shown that there are always created spirits who do or can know the truth of the matter, and that there is reason to think that things that make no difference from the point of view of the persons themselves will do so only temporarily. (NE, 2, 27, § 29)

\*

This is a comprehensive presentation of Leibniz's explanations which he advances in his doctrine. The construction of this line of argumentation, even though it is not without tensions, is coherent in its main strand. One unquestionable advantage of Leibniz's theory is the balance between metaphysical and psychological dimensions of personal identity. In this way, personal identity over time gains a strong metaphysical basis, which, however, does not undermine the prominent role that is played by self-consciousness and memory

in the structure of a human being. The belief that personal identity is based, on the one hand, on the continuity of substance (substantial “I”) and, on the other, on the continuity of self-consciousness and memory opens – so it seems – the way to a more comprehensive insight into the conditions of personal identity.

## Acknowledgment

The article was supported by a research grant from the National Science Centre, Poland (UMO2012/07/B/HS1/01959).

## Bibliography

### Primary Sources

- Erdmann Leibniz, Gottfried William. 1840. *Opera philosophica omnia*. Edited by Johann Eduard Erdmann. Berlin: Scientia Verlag Aalen.
- GP Leibniz, Gottfried William. 1875–1890. *Die philosophischen Schriften von G. W. Leibniz*. Edited by Carl Immanuel Gerhardt, vol. 1–7. Berlin: Weidmann.
- Gr Leibniz, Gottfried William 1948. *Textes Inédits*. Edited by Gaston Grua, vol. 1–2. Paris: Presses Universitaires de France.
- L Leibniz, Gottfried William. 1969. *Philosophical Papers and Letters*. Translated and edited by Leroy E. Loemker. Dordrecht: D. Reidel.
- NE Leibniz, Gottfried William. 1996. *New Essays Concerning Human Understanding*. Translated and edited by Peter Remnant and Jonathan Bennett, 2<sup>nd</sup> ed. Cambridge: Cambridge University Press.
- AG Leibniz, Gottfried William. 1989. *Philosophical Essays*. Translated and edited by Roger Ariew and Daniel Garber. Indianapolis and Cambridge: Hackett Publishing Co.
- Theodicy Leibniz, Gottfried William. 2007. *Theodicy, Essays on the Goodness of God, the Freedom of Man and the Origin of Evil*. Translated by Eveleen M. Huggard, edited by Austin Farrer. Oxford: Biblio Bazaar.

### Secondary Sources

- Bobro, Marc Elliot. 2004. *Self and Substance in Leibniz*. New York: Springer-Verlag.
- Curley, Edwin. 1982. “Leibniz on Locke on Personal Identity.” In *Leibniz: Critical and Interpretative Essays*, edited by Michael Hooker, 302–26. Minneapolis: University of Minnesota Press.
- Graganiec, Mariusz. 2016. “O tożsamości osobowej. Leibnizjańskie wątki koncepcji Rodericka M. Chisholma.” *Przegląd Filozoficzny – Nowa Seria* 100(4): 405–19.

- Jolley, Nicholas. 1984. *Leibniz and Locke: a study of New Essays Concerning Human Understanding*. Oxford: Clarendon Press.
- Mijuskovic, Ben Lazari. 1975. "Locke and Leibniz on Personal Identity." *The Southern Journal of Philosophy* 13(2): 205–14.
- Noonan, Harold. 1989. *Personal Identity*. London: Routledge.
- Scheffler, Samuel. 1976. "Leibniz on Personal Identity and Moral Personality", *Studia Leibnitiana* 8(8): 219–40.
- Mates, Benson. 1986. *The Philosophy of Leibniz: Metaphysics and Language*. New York, Oxford: Oxford University Press.
- Thiel, Udo. 1998. "Personal Identity." In *The Cambridge History of Seventeenth-century Philosophy*, edited by Daniel Garber and Michael Ayers, 868–913. Cambridge: Cambridge University Press.
- Vailati, Ezio. 1985. "Leibniz's Theory of Personal Identity in the New Essays." *Studia Leibnitiana* 17(1): 36–43.
- Wilson, Margaret. 1999. "Leibniz: Self-Consciousness and Immortality in the Paris Notes and After." In Margaret Wilson, *Ideas and Mechanism*, 373–87. Princeton: Princeton University Press.

# Diderot and Materialist Theories of Self

“I am me because my little dog knows me but perhaps he does not and if he did I would not be me.”

(Gertrude Stein, “Basket a story interlude,”  
in Stein 1936/1995, 198)

## Introduction

A materialist theory of self must be able to respond to certain objections, explain certain phenomena and reach a certain degree of articulation. The same applies for specifically early modern versions thereof, which shall be my concern in what follows. Our familiarity with the diverse forms of early modern materialism has grown a great deal in recent decades, marking a considerable advance over older (if still regrettably common at times) views of materialism as inherently mechanistic, fixated on the idea of the body as machine, or denying basic features of embodiment. But what of the self? I shall discuss the materialist treatment of the self, and overall the cluster of problems concerning selfhood, individuality and personal identity in various authors, but most centrally in Diderot. My analysis is neither a standard internalist reconstruction of a problem in Diderot, with passing mention of other period authors, nor an intellectual history-type survey of a problem in the period, with discussion of as many authors as possible. It is, as the title indicates, a reflection on Diderot *and* materialist theories of the self. That is, I aim to reconstruct a problem, and will suggest that Diderot puts forth one of the more significant and original versions of a materialist theory of the self – but one which, of course, appropriates elements from other authors.

The self was often seen as simply a part of the classic ‘matter and mind’ problem. Thus the salonist Suzanne Necker (later mother of Mme de Staël) reprises classic Cartesian points but to speak of the self: “half of a *self* is a contradictory absurdity, while a portion of matter that cannot be divided is also a contradiction: how can mind and matter not be different substances?” (Necker 1798, III, 88). One should note that this shift to the problem of the self presents a particular kind of conceptual challenge. Why should the materialist approach to the self be particularly challenging? Because the latter belongs to a time-honoured family of philosophical intuitions which are perennially presented as light years removed from the world of materialism. From Augustine (*Confessions*, X, 16, 25) to

Descartes and on to Paul Ricoeur (1992), or from Kant and Schelling onto Husserl and Heidegger (but also Wittgenstein, Anscombe, Chisholm, Nagel, etc.), we are told in endlessly varied ways that the self is not, to borrow Wallace Stevens' elegant line, "composed of the external world":<sup>1</sup> that the self is not of the material world, whether this has to do with its lack of divisibility, its temporal essence, the inner sense, grammatical properties of the first person, or other "facts".

Contrasting with such views (or intuitions, which is often what they are), I point to the existence of an early modern materialist discussion of self – an intellectual "tradition", even if it lacks direct transmission or continuity. It is early modern inasmuch as it extends, in the authors I focus on here, from Spinoza and Locke to Diderot. Indeed, I do not believe that the attempt to combine thoroughgoing materialism and a concept of self is somehow a "timeless" feature of materialist thought, and it may well be the case (although I make no such metahistorical claims on my own account) that concern with the self is a post-Cartesian development, in the sense of the Augustinian elements in Descartes, or even Luther on some readings (see Menn 1998 on the former and Schürmann 2003 on the latter).

This materialist approach to the self can take (at least) three forms, which occur independently of one another (e.g. in Spinoza or La Mettrie) but which can also be combined, as they are in admittedly programmatic form in Diderot. These are: externalism as a metaphysical position (§ 2), the biologization of individuality, i.e. a justification of individuality in biological terms (§ 3), and the equation of brain and self, in a reductionist approach to the problem of personal identity (§ 4), although ultimately 2 and 3 are the basic "planks" of the theory. In conclusion (§ 5) I suggest that rather than being "blind to the world of internal life" as was often claimed of materialism, there can be something like a materialist theory of self, notably but not exclusively as sketched in Diderot. Differently put, rather than a whole-scale elimination of the mental, the early modern materialist approach could also be a "naturalization" of the mental – an inscription of mental life in the broader natural world, which does not make it disappear as if by waving a wand (a separate issue is how early modern materialism might relate, or not relate to the emergence of psychology as a science).

## Externalism

Discussions of "person", "self", "experience", even when they bring in an embodied, material dimension, frequently appeal to a *first-person* concept of experience. This is usually opposed to a *third-person* view, typically presented as the point of view of the natural scientist with her measuring instruments. Many philosophers hold that we will *never know what it is like* to have someone else's first-person experience. There is something here like an opposi-

---

1 To be clear, Stevens's line goes the other way: "the soul, he said, is composed of the external world." (Stevens, "Anecdote of men by the thousand", in Stevens 1997, 41).

tion between the internal and external; between internalists and externalists, in the sense I will elaborate on below (namely, about epistemic content and indeed about the status of minds in the physical world, not as semantic theories). I will defend a version of externalism in the following section.

By externalism I do not mean a semantic theory or a social theory of mind (at least two of the other senses of the term)<sup>2</sup> but rather the position according to which mental states lack any inaccessible, “first-person” dimension; any such dimension would be either explainable in external terms or traceable to processes in the agent (the nature of which need not be specified, as indeed the Identity Theorists of the 1960s–1970s did not, but we could think of them as brain processes), which produce a “feeling” of interiority. The externalist does not need to deny that a person has experiences, but she does not hold that experience is thereby inaccessible, radically private, incommensurate with the rest of the physical universe. If the internalist holds that “States, or experiences [...] owe their identity as particulars to the identity of the person whose states or experiences they are”, as in Cudworth’s conception of the self as *to hegemonikon* or as defined by *sui potestas*, endlessly echoing itself (Cudworth 1996, X, 178), the externalist holds that “no fact is only accessible to a single person”<sup>3</sup> and deplores, as Diderot does in § X of his 1753 *Pensées sur l’interprétation de la nature*, that it is easier to consult oneself than to consult Nature.

The externalist will hold that any sense of unity, any foundational dimension of selfhood, in fact comes from *outside*. Materialism implies externalism but externalism does not imply or entail materialism (a vision of the mind as social, including as behaviourally constituted in a world of activity, is not committed to a materialist metaphysics).<sup>4</sup>

One can also see the distinction between internalism and externalism in the difference, familiar to scholars, between the Cartesian *cogito* and the Spinozist *homo cogitat* (*Ethics* IIa2). That “*homo cogitat*” is not a foundational property of a first person; the self, and its key property, thinking, is not *foundational*. To be a thinking subject is simply to belong to the universe of causal relations, to be a particular intersection within it (compare Renz

---

2 Notably in Putnam’s sense. Tyler Burge’s “social externalism” (e.g. Burge 1979) takes an ontological step further but is still about “mental contents”, whereas the “Spinozist” externalism I am suggesting here is straightforwardly ontological (if any such thing is straightforward), although it shares the suspicion (or deflationary attitude) towards any a priori, privileged self-access. Thanks to Ville Paukkonen for making me clarify this point.

3 Strawson 1959, 97 and Dretske 1995, 65 (although of course some would accept this claim and nevertheless argue that some objects of knowledge are not facts – thanks to an anonymous reviewer for this point).

4 It is possible for there to be a materialist who is not also an externalist, in a rather forced thought experiment in which there only exists one atom. But from Lucretius to Hobbes, the anonymous *Theophrastus redivivus* of 1659, Collins, La Mettrie, d’Holbach, Priestley, Cabanis, Büchner and Vogt, Dewey, the Spinozist social psychologist Lev Vygotsky, Quine in some moods, D.M. Armstrong, David Lewis and present-day physicalism, materialists are externalists (names missing from this list are not non-externalist materialists, but rather materialists like Toland or J. J. C. Smart who do not offer a position on the matter). On externalism in a non-materialist context as a social theory of mind in Locke, see Lenz 2013, and for a different, but complementary perspective, Wolfe 2010.



2017). In Spinoza’s memorable phrase, “The order and the connection of ideas is the same as the order and the connection of things.”<sup>5</sup> For the externalist, no fact, datum or *vécu* belongs to a private, off-limits zone, for what is first is not the thinker but the web of relations to which thought belongs. Of course, Spinoza doesn’t content himself with this static vision of a grid of relations; he emphasizes that any such particular “individuated” entity strives to persevere in existence, as the finite mode it is. I cannot improve on Morfino’s summary:

[F]or Spinoza the individual is neither substance nor subject [but...] is a relation between an outside and an inside constituted by this very relation (there is no absolute interiority of the *cogito* opposed to the absolute exteriority of a world). This relation constitutes the essence of the individual, comprised of its own existence-power. [...] It is a variable power, precisely because the constitutive relation between inner and outer is unstable, not established. The passions are not, therefore, the property of an already given human nature, but they are relations constituting the human individual; *their locus is not interiority, but the space between individuals*. (Morfino 2006, 118; trans. modified, emphasis mine)

The externalist has a *relational* definition of what it is to be an individual, as a particular portion of a given, causally closed space-time, of a state of relations which constitutes a given individual – an oak tree, a stag beetle, Mutlu the cat – *qua* that which resists decomposition (a “conatus ad existendum”). Of course, to claim that Spinoza defines the individual as a relation, or gives ontological primacy to relation, may seem to run counter to the obvious fact that Spinoza thinks the individual is defined by its own conatus, its own essence (EIIIp9s: the conatus is our essence). Yet the relational view has in favour of it equally core Spinozist definitions: our body needs a great number of other bodies to survive (EIIp13, 4<sup>th</sup> postulate), just as our mind would be imperfect if it only took itself as an object (EIVp18s). In addition, bodies form a single body or individual when their movements are related to one another (or when they “communicate” according to a precise ratio or relation: EIIp13d). In sum, we are defined by a certain ratio, proportion or relation of motion and rest (*ratio motus et quietis*).<sup>6</sup>

For the externalist, an experience, a desire, or a belief do not belong *de jure* to a constitutive subject, but rather *de facto*, to a subject which *they* constitute. As Dewey put it,

5 E IIp7; see also IIIp2s. This is an ontological assertion of the primacy of relations; some readers, including Vygotsky, also viewed it as expressing the non-independence of mind and brain with regard to this world. (Vygotsky 1972, 362–82).

6 Beyond the comment to Jarig Jelles, Spinoza can handle individuality structurally, as a particular ratio of motion and rest: E IIp13s (the physics), esp. lemmas 1 and 7s.; *Short Treatise*, appendix, II.14 – leading, however, to troubles such as the case of the Spanish poet (E IVp39s). At the structural level of ratios, he is the same person, *certa quadam ratione*; at the level of his mind, he is not. See also Toto 2015, 65.

We live from birth to death in a world of persons and things which is in large measure what it is because of what has been done and transmitted from previous human activities. When this fact is ignored, experience is treated as if it were something which goes on exclusively inside an individual's body and mind. It ought not to be necessary to say that experience does not occur in a vacuum. There are sources outside an individual which give rise to experience. (Dewey 1938/1963, 39)

Indeed, the subject is constituted by her progressive filtering (and filtering out) of the world, which also serves as an argument against scepticism, according to the idea that the senses are made *for* x. This sensory filtering is described in Diderot's important, but at the time unpublished *Rêve de D'Alembert* (1769) as constitutive of our individuality: no one's sensory make-up is identical to anyone else's sensory make-up. "The animal is a unified whole" for Diderot, both because of its specific physiological constitution (*organisation*) and specifically because of what he calls its organic *continuity*, as distinct from the mere *contiguity* of parts.<sup>7</sup>

The limits of my sensory system are also my limits as an individual, in the sense that however much all of matter may be living matter, I cannot sense what is happening on Saturn, for between me and this planet "there are only contiguous bodies, instead of continuity" (*Rêve de D'Alembert*, in Diderot 1975–, XVII, 142). Elsewhere, in the *Éléments de physiologie* (hereafter EP), Diderot puts it this way: "if external sensations [...] and inner sensations were equally intimate to me, *everything* would be me, and I would be *everything*" (Diderot 1975–, XVII, 460). I don't perceive the cosmos directly (my perceptual apparatus acts as a filter); if I did, the barriers of my self would somehow be the barriers of the world. For sensation (perception, experience) are both real and constitutive of self here.

The self is constituted from without, and the sensory part of this process entails that no two subjects will perceive the same object in the same fashion. This is the properly materialist way of accepting that someone's life history, including the larger-scale evolutionary history, is constitutive of their being. Notice that we have a criterion of personal identity here: "For any organism *x* and any *y*,  $x = y$  if and only if *x*'s life is *y*'s life" (Olson 1997, 138). And since externalism does not mean that my self is equal to the universe as a whole, we can see something of a biological emphasis being smuggled in here. If I am not defined by a free, unconditioned inner space of interiority, but by a multitude of "petites perceptions" (often interpreted in determinist and materialist terms in the early eighteenth century, e.g. by Anthony Collins in his *Inquiry Concerning Human Liberty* of 1717; see Wolfe 2007) crisscrossing in my mental life, by my physiological constitution, by "the blood which flows

---

<sup>7</sup> Diderot, *Éléments de physiologie* (hereafter *ÉP*), in Diderot 1975–, XVII, 335; *Rêve de D'Alembert*, in Diderot 1975–, XVII, 140, 142. Diderot always insists on the specificity of each individual's *organisation*.

in my veins”, as La Mettrie would have it,<sup>8</sup> then we have gradually shifted from externalism *per se* to a *biologization of individuality*.

## The Organic Self

There is nothing novel or particularly radical about philosophy turning to the biological world to obtain its “best definition” of what an individual substance is; think of Aristotle, who tended to use actual organisms as paradigm cases of individual substances, or in contemporary parlance, “paradigmatic individuals”.<sup>9</sup> But it is a further step to say that the traits associated with our interiority are themselves biological in nature – whether it be the “inner sense”, intentionality, the synthetic unity of apperception, consciousness, and so on. Indeed, one author, Nietzsche, warned in the late nineteenth century against committing a sort of category mistake and confusing the self with the “feeling of organic unity” (“Das Ich – nicht zu verwechseln mit dem organischen Einheitsgeföhle”: fragment from Spring–Fall 1881 = M III, 11(14), in Nietzsche 1973). I am interested in the narrower class of thinkers who explicitly disobey the Nietzschean warning not to confuse the self with the feeling of organic unity, or in more general terms, who think that *facts about selves*, including experiential ones, *might turn out to be biological facts*, i.e. that “personal” facts are actually “organismic” facts, and as such (unless biology should be restricted to genetics or molecular biology), biological facts. Of course, even in this narrower class we can find the argument running in two contrasting directions: either

— a *reductionist* direction, in which the thinker will retain whichever experiential, existential or phenomenal properties can be successfully preserved after a reduction to the biological facts

or

— a *holist* direction, in which there is a “transfer” of subjective properties onto biological entities, usually the “organism” (which is one major reason for the bad

---

8 For La Mettrie, as he details in his *Discours sur le Bonheur*, each of us, the criminal and the honest man, are in pursuit of our own good – happiness, particularly understood as pleasure; whether I am virtuous or vicious depends “on my blood”: surely something individual, but equally surely, not the sort of individuality most philosophers of personhood would be happy with.

9 Aristotle often insists that animals (not artefacts) are the paramount case of individual substances (*Metaph. Z.7* 1032b19); Aristotelian scholarship since at least Montgomery Furth’s tour de force work (ultimately presented in Furth 1988) has spoken of animals as “paradigmatic substances”: “Animals, in Aristotle’s view, are paradigm instances of substance-being” (Kosman 1987, 360). For the notion of individual substance in contemporary philosophy of biology, see Hull 1992, 182 and Richards 2010, 164–5 for discussion.

reputation of the concept of organism in some circles, as it is taken to be a kind of “last gasp” of Romantic subjectivism within biological theory).

Both of these are naturalistic, but the former squares more easily with most understandings of naturalism (admittedly a rather loose term), while the latter, because it will insist on retaining a modicum of subjective language (from self and selfhood to “inner states”, “experience” and the like), can be deemed unnaturalistic on some accounts. Now, the more reductionist version of a biological theory of self will equate “self” with a set of bodily or even cognitive processes or states, to be specified (and this place-holder quality raises further questions I do not address here, such as: is it an evolutionary account? cerebral? neurobiological? etc.), which is not the same as the eliminativist view according to which no such thing as the self exists.

Why is the above reductionist option not the same as eliminativism? To take a classic example from a self-proclaimed early modern materialist, La Mettrie: when he writes that “The soul is just a pointless term of which we have no idea and which a good mind should only use to refer to that part of us which thinks” (La Mettrie 1987, I, 98), is this reductionist or eliminativist? Contemporary terminology relies on the distinction between reductionism and eliminativism, both of which have a respectable materialist pedigree. In the above case (the existence of the soul), eliminativism holds that the soul and all of its properties that have been described and argued over from, say, antiquity and Scholasticism through Swedenborg *does not exist* and indeed *none of these properties are real*; thus, what *is* real would be the brain, or the heart, or the stomach, and so on. Reductionism holds that the soul (to stay with the same example) is indeed not something that exists in any traditional sense; but notice that when La Mettrie says above that we really should only use the word to refer to “that part of us which thinks”, he is not saying mental faculties do not exist but that we need to rethink what their “seat” is, where they come from, and the extent to which they are independent from the rest of bodily processes, or not. However, he is not suggesting a weaker thesis, which would be that soul/mind might be autonomous in some sense but could be “defined in terms of” bodily processes. The materialist theories of self discussed here share a commitment to reductionism, but not to eliminativism (although the extent to which this distinction is clearly applicable to the texts at hand is unclear).<sup>10</sup>

I shall take Diderot as my major example of the biologization of individuality. For Diderot, materialism definitely implies a degree of reduction – a deflationary or destructive impulse to trace back, as he writes to Damilaville, “our most sublime feelings and our purest tenderness” to “a bit of testicle” (Nov. 1760 letter, in Diderot 1955–1970, III, 216). But this is not a reduction of human or animal action or personhood to the action and necessitation

---

10 Not just in the early modern context, but also in contemporary philosophy of mind, some authors are not clearly reductionist or eliminativist despite being naturalists, such as Dennett who views the self as a fictional entity, but one which is useful in evolutionary terms (Sturm 2007, 173). A significant precursor of Dennett’s here is, of course, Dewey.

of falling stones or clockwork. It is a reduction to the animal, so to speak – as when Diderot commented critically on Hemsterhuis' 1772 *Lettre sur l'homme*, “wherever I read *soul* I replace it with *man* or *animal*” (Diderot 1975–, XXIV, 340). It retains an embodied focus, so that, e.g. properties of the soul are explained in terms of properties of the body (*ÉP*, Diderot 1975–, XVII, 334–335), not of fundamental physics. In the language of theory reduction, we could say that for Diderot, the reducing theory is biology, not physics (there was no physics to speak of, and more importantly, he felt that the cluster of theories later to be termed biology, and then referred to as “natural history” as a catch-all term, which indeed could mean the science of life in general, was the richest).

In the *Rêve de D'Alembert*, the character D'Alembert challenges the character Diderot to account for the self. Diderot has more or less successfully defended the concept of a living, sensing and thinking matter, but D'Alembert queries: “Could you tell me about the existence of a sentient being in relation to itself?”, that is, about the self-awareness of a sentient being. Diderot speaks in Lockean terms of memory as the basis for our self, with the materialist twist that memory itself is the product of our physiology (*organisation*). But later in this work the character Mlle de Lespinasse states how obvious it is to her that she is herself: “it seems to me that there is no need of such verbiage to know that I am me, I have always been me, and I will never be any other” (Diderot 1975–, XVII, 134).

Diderot's materialist reply is that the self is itself the result of a construction of smaller elements – parcels of living matter. An organism is formed by adjunction of living points or animalcules, by purely material processes: “A hundred, a thousand times, I have seen the shift from inert matter to active sensitivity, to the soul, to thought, to reasoning – without any other agent or intermediary than material agents or intermediaries.” (*Observations sur Hemsterhuis*, in Diderot 1975–, 277). In this shift from inert matter to sensing, living matter, how do I *feel* that I am myself? For Diderot, the answer is: in and through my central nervous system – which is both *myself* and a guarantor of my relation to the rest of the material world in a constant process of exchange.

Diderot is one of the first materialists to explicitly take note of the “fact” that organisms are in part defined by their sense of unity, a unity he describes in the language of unified causality:

without regard for the sum of elements of which I am composed, I am one, and a cause only has one effect; I have always been one single cause [*une cause une*], thus I have never had more than one effect to produce; my duration is thus nothing more than a succession of necessary effects. (*Jacques le fataliste*, in Diderot 1975–, XXIII, 190, 28)

In that sense, I cannot “do otherwise than myself” or “be anything other than myself” (*ibid.*). Diderot does not provide an extended philosophical commentary on this “single cause” or unified selfhood, but it is clearly a recurrent concern in his work, whether he is faced with its denial in other, more reductionist projects or its defence in “dualist” or otherwise implausible projects which neglect, for example, determinism.

Diderot's articulation of an embodied materialism – not synonymous with “physicalism” – can help itself to some of the key features of selfhood, individuality and identity, which anti-materialists from More and Cudworth to Reid and Husserl insisted could not be present in a materialist analysis.<sup>11</sup> Commentators often overlook Diderot's critique of Helvétius' *De L'Homme* (1773), which precisely focuses on the latter's excessively “mechanistic” picture of behaviour as subject to standardized rules of social conditioning. Now, Diderot disagrees with Helvétius' “social determinism” of operant conditioning, but unlike Cudworth, Clarke, Reid or Madame Necker, he does not do so in the name of an unconditioned, uncaused or otherwise “extra-territorial” self (Wolfe 2007). He finds Helvétius' programme to be not only dangerous but condemned to fail, at the very least because of the irreducible “organic” or “psycho-physiological” specificities of each individual. But within that organic individuality, there is no homuncular self.

In that sense, the judgment, found in a study of Diderot, that “Materialism as a working philosophy, used as a tool in the scientific investigation of the material universe, is appropriate and highly effective. Intended for the objective analysis and description of the world of externals, it yields disastrous results when applied to the inner, subjective world of human nature, human thought, and human emotions,”<sup>12</sup> is at best the wielding of a very blunt explanatory instrument, and at worst, a projection of a personal valuative decision onto seventeenth- and eighteenth-century texts. Both La Mettrie and Diderot, and most of their critics in the eighteenth century, would have been surprised to hear that materialism was an effective tool for science and for handling “the world of externals”, but not for the inner life. Perhaps Charles Bonnet would have been satisfied by this distinction, for he believed that his wholly naturalistic, causal, mechanistic analysis of the neurophysiological bases and correlates of psychological association was at the same time, non-materialist. And what is this inner, subjective world such that it should be left untouched by the materialist, who knows only “the world of externals”? As the French neuroscientist Marc Jeannerod wrote, it is dangerous to leave subjectivity to the philosophers...<sup>13</sup>

---

11 More 1653, XI; Cudworth 1996, as cited earlier; Reid 1788/2010, 4.4, and discussion in Yaffe 2004; Husserl 1989, § 32.

12 Hill 1968, 90. I have developed my criticisms of this view of materialism a bit further (including with respect to ethics) in Wolfe 2016.

13 Jeannerod 1983, 121. Jeannerod had in mind a kind of intellectual abandonment in which the life of the mind, the preserve of philosophers, is taken to be “subjective” and “qualitative”, while cerebral life per se would be “external” and “quantitative”. This is among other things, as he observes, an extremely impoverished vision of causality.

# Self and Personal Identity

If the biologization of individuality seems to enable the materialist to do justice to some core features of selfhood (on the condition that she is not a strict physicalist, in which case facts about the self would be declassified from any material standing, and relegated to qualia, folk psychology, etc.), the same cannot be said, or at least not as easily, of externalism. Thus a “qualitative” argument against externalism (which is, however, quite compatible with biological theories of individuality) will declare that there is *something* that it is like to be me, a special relation, which cannot be grasped from outside, and *a fortiori* by the scientific, ‘third-person’ perspective. The world of relations seems to “drown” individuality: this seems to have been Montesquieu’s reaction to Spinoza, which I cite not least because of its vivid turn of phrase: he felt that Spinoza “deprived him of everything personal,” so he could no longer “find that self in which I was so interested”; “why glory? why shame? [...] in the universality of substance, both the lion and the insect have come and gone indistinguishably, both Charlemagne and Chilpéric” (Citton 2006, citing Montesquieu at 77).

But if materialism is granted, should selfhood be located (a) in a set of relations, as a structurally defined feature, a “ratio of motion and rest” in Spinozist terms (as in *Ethics* IIp13s), (b) in an actualized, temporal, finite biological entity – with additional individuating features to be specified involving its homeostatic equilibrium, its immune system, and so forth, or (c) purely in processes, such as Locke’s continuity of consciousness over time?

Recall that Locke’s celebrated theory of personal identity was in large part intended to avoid having to locate the latter in a merely material substance: “[those] who place Thought in a purely material, animal Constitution, void of an immaterial Substance” plainly “conceive personal Identity preserved in something else than Identity of Substance; as animal Identity is preserved in Identity of Life, and not of Substance” (Locke 1975, II.xxvii.12.). In addition to this “identity of Life,” humans have a form of reflexive self-consciousness, a type of “privileged access” to ourselves in our ability to remember our past – despite problems such as potentially fabricated memories – which we do not have in relation to others, including the narratives of others.

We are dealing here with memory, a type of privileged access crucial enough for it to be constitutive of personal identity itself. Yet Locke doesn’t hold that memory *per se* is the guarantor of personal identity. This is what I termed a “processual” definition of selfhood: it explicitly aims to replace any substantial definition – including, of course a materialist definition. Of course, Locke is frequently agnostic about tensions between immaterialism and materialism, but in the present context he seems to lean in one direction: “the more probable opinion is that this consciousness is annexed to, and the affection of one individual immaterial Substance.”<sup>14</sup>

---

14 Locke 1975, II.xxvii.25. However, two points allow for differing interpretations: the first, the extent to which Locke is “agnostic” about materialism, and the second, the extent to which Diderot’s materialist



Is a materialist approach to personal identity instantly invalidated, or at least weakened, by Locke's anti-substantialist theory? Yes, if it meant understanding what a self or individual is (granted, these are not identical terms!) in strictly aggregative terms. To be clear, concepts of selfhood and of individuality are often run into each other in the texts of the period, as has been observed by Udo Thiel in one of the best studies of the topic (Thiel 2006, 296). But Locke's important insights are not fatal to a more organismic (and thus also relational) concept of self. That sentence is a feature of advanced organisms is taken by Diderot as an empirical fact (deriving from experiments such as Haller's on the nervous system: cf. Wolfe 2014; Duchesneau 2017). Granted, for a "Cartesian" or "Kantian", empirical facts are certainly not relevant to a decision about the nature of the mind, but that is a problem beyond the scope of this paper.

Recall Diderot's distinction between merely spatial and mechanical contiguity, and properly organic, indeed organismic, continuity: the latter concept includes an existential, processual, temporal dimension, in the sense that an organism is not just a "snapshot" of an organism. To cite Olson again, "For any organism  $x$  and any  $y$ ,  $x = y$  if and only if  $x$ 's life is  $y$ 's life." And the sophisticated materialist theorist of personal identity, not least a biologically inspired materialist, should not be unaware of the simple fact that the cells in our bodies change over time (an example which Locke thought was fatal to a naïve substantialist-materialist theory of personal identity). As Diderot himself reflects in *Rêve*: "through all the vicissitudes I experience in the course of my duration, given that I may not possess a single one of the molecules I was composed of at birth, how did I remain myself to others and to myself?" (Diderot 1975–, XVII, 163). Here the Spinozist point that what it is to be me is not so much a fixed set of material parts, but rather a *ratio*, is applicable. Think also of the case of our immune system (Pradeu 2012), which is neither reducible to a "thing" located at one fixed point in time and space, nor a *cosa mentale* which the biologically nourished materialist can say nothing about.

Yet the structural answer (which corresponds in more detail to what I have called "externalism") does not exhaust the materialist treatment of personal identity. In fact, Locke's emphasis on memory can be integrated therein, despite the seeming paradox (since it was intended to reject the material substantiality of the self). This integration is notably possible because of the shift in our understanding of memory as itself a cerebral function. That is, Locke rejects material criteria for personal identity and asserts the criterion of memory; but we would say today that the mechanisms of memory are cerebral functions! As Ludwig

---

theory of memory can nevertheless be said to be Lockean or not (if personal identity is constituted by the processual unification of thoughts and actions through memory, one can be "neutral" as regards the substrate of this process: it is material in Diderot, but Locke would not have to disagree with this). Thiel notes that some early critics of Locke even thought that Locke deliberately accounted for personal identity in terms of consciousness and memory, and not in terms of the same substance, so as to open the door for a materialist account of the mind: thus, Isaac Watts suspected that Locke "is so very solicitous to make the *same Substance* unnecessary to *Personal Identity*, that so he may maintain his supposed Possibility of Matter being made capable of Thinking" (Watts, *Philosophical Essays*, 3<sup>rd</sup> ed 1742, 302, cit. Thiel 2011, 150). I thank Udo Thiel for discussion about this point.



Büchner put it, “memory is merely the product of material combinations” (Büchner 1870, 131). Indeed, Diderot himself described memory as a “corporeal quality” (*EP*, in Diderot 1975–, XVII, 335) but also appeals to it in very Lockean ways, for instance when he criticizes Hemsterhuis’ version of a traditional immaterialist concept of personhood, stressing that without the memory attached to a series of actions, the individual, moving from sleep to wakefulness and back again, would barely be able to take note of her own existence. At the same time, this apparently “processual” rather than “substantial” concept is also integrated in Diderot’s conception of what I called above “the organic self”, as when he asserts that “the history of the life and the self of each animal is composed of the memory of its successive impressions” (*Rêve*, DPV XVII, 155). The structural here has become the corporeal, and/or the cerebral.<sup>15</sup>

## Conclusion

The materialist theory of self need not be blind to or dismissive of all features of interiority. While it is necessarily deflationary or reductionist towards selfhood qua interiority, certainly as something foundational (the early modern materialist could very well have said “You are *not* authoritative about what is happening in you, but only about what *seems* to be happening in you.” Dennett 1991, 96), the theory can, notably, integrate degrees of embodied selfhood, qua biological individuality, given that it is not an outright physicalism (although depending how much a given thinker builds into their physics, like Hobbes and his “small beginnings of motion” which account for both physical and mental forms of striving,<sup>16</sup> one can arguably go some distance towards an account of volitional and other parts of mental life on such a physicalist basis...). Instead of denying the existence of interiority, the materialist should try and locate it within the physical world, within the overall framework of explanation (as Spinoza did). But since this materialism is not strictly a physicalism but can appeal to biological information, it offers plenty of ways to understand individuality, selfhood or agency – as in the “immunological self” (Pradeu 2012). The articulation of externalism and the specifically biological or embodied dimensions of certain forms of materialism can be expressed in a mantra of social psychology (here,

---

15 In this sense, Diderot may be an interesting “problem case” for the opposition between scholars such as Mijuskovic and Thiel, for Thiel rejects Mijuskovic’s claim that materialist theories seek to establish “personal identity on a model of bodily identity” (Mijuskovic 1974, 105) and emphasizes instead the Lockean dimension, according to which materialist theories of personal identity accept arguments against material-substantial continuity (Thiel 1998, 69). In this paper, Thiel suggested that “French materialist *philosophes* do not concern themselves very much with the special problem of personal identity” (Thiel 1998, 63n.); perhaps Diderot’s *Rêve de D’Alembert* merits a revision of this claim. (Thiel updated his views in his 2015 paper.)

16 Hobbes 1976, chapter 33, § 2, 407; *De Cive*, IX, 9, in Hobbes 1992, II, 120, and Hobbes 1994, XXI, 1.

prominent social psychologist Roy Baumeister): “Everywhere in the world, self starts with body.” (Baumeister 1999, 2).

The point is not that the materialist theory of self, for instance in Diderot’s version, encompasses all the positive features of all other theories of self without any of their negative features, but that classic oppositions between a world of agency, value, intentional states and privacy, and a “merely spatial” and/or mechanical and by extension somehow dehumanized world, need serious revision. This overlaps with a related problem concerning early modern materialism, when it is understood as somehow necessarily mechanistic (Wolfe 2017).

The theory as I have reconstructed it essentially comprises a “relational”, externalist metaphysics and a biological vision of individuality. The advantage of the biological perspective is that it preserves a certain realism; the power but also the danger of externalism as an ontology of relations, and of the reduction of personal identity, is that they lose trace of any existence of the self (as Spinoza was often reproached: the “selfhood” of one finite mode among others does not seem like the most appealing defence of the self). But this advantage – unless one has a kind of transcendental criterion with which to automatically reject any confusion between the self and the “feeling of organic unity” – brings with it the danger of “biologism”, and of a metaphysics of the organism, in which the fact that certain features of selfhood seem to be present in (some parts of) the biological world allow for a kind of a re-transcendentalization, notably of “the organism”. Hence the materialist theory of the self is a mobile (and modular) set of concepts, with its advantages and its disadvantages, its diversity and its limitations. Future histories or philosophical *survol*s of the self might consider it worthy of inclusion.

## Acknowledgments

I wish to thank Udo Thiel for his comments on an earlier version of this paper, as well as Judit Szalai and an anonymous reviewer for their further suggestions. Final work on this essay was carried out as part of a project that has received funding from the European Union’s Horizon 2020 Research and Innovation Programme (GA n. 725883 ERC-Early-ModernCosmology).

## Bibliography

### Primary Sources

Büchner, Ludwig. 1870. *Force and matter: empirico-philosophical studies, intelligibly rendered, with an additional introduction expressly written for the English edition*. London: Trübner.

- Cudworth, Ralph. 1996. *A Treatise concerning Eternal and Immutable Morality*, reprinted with the *Treatise on Free-Will*. Edited by S. Hutton. Cambridge: Cambridge University Press.
- Diderot, Denis. 1955–1970. *Correspondance*. Edited by G. Roth, 9 vols. Paris: Éditions de Minuit.
- Diderot, Denis. 1975–. *Oeuvres complètes*. Edited by H. Dieckmann et al. Paris: Hermann.
- Hobbes, Thomas. 1976. *Thomas White's "De Mundo" Examined*. Translated by H. W. Jones. London: Bradford University Press.
- Hobbes, Thomas. 1992. *English Works*. Edited by W. Molesworth, 12 vols. London, 1839–1845; reprint. London: Routledge/Thoemmes.
- Hobbes, Thomas. 1994. *Leviathan*. Edited by E. Curley, with selected Latin variants. Indianapolis: Hackett.
- La Mettrie, Julien Offray de. 1987. *Œuvres philosophiques (1751)*. Edited by F. Markovits, 2 vols. Paris: Fayard–“Corpus”.
- Locke, John. 1975. *An Essay Concerning Human Understanding*. Edited by P. Nidditch. Oxford: Oxford University Press.
- More, Henry. 1653. *An antidote against atheisme, or, An appeal to the natural faculties of the minde of man, whether there be not a God*. London: Roger Daniel.
- Necker, Suzanne. 1798. *Mélanges, extrait de ses manuscrits*. Paris: Charles Pougens.
- Nietzsche, Friedrich. 1973. *Werke. Kritische Gesamtausgabe*, 5. Abteilung – 2. Band, *Idyllen aus Messina. Die fröhliche Wissenschaft. Nachgelassene Fragmente Frühjahr 1881 bis Sommer 1882*. Edited by Giorgio Colli and Mazzino Montinari. Berlin–New York: Walter de Gruyter. <http://www.nietzschesource.org/#eKGWB/NF-1881,11>
- Reid, Thomas. 2010. *Essays on the Active Powers of Man (1788)*. Edited by Knud Haakonssen and James A. Harris. University Park, PA: Pennsylvania State University Press.

## Secondary Sources

- Baumeister, Roy F. 1999. “The nature and structure of the self: An overview.” In *The Self in social psychology*, edited by R. Baumeister, 1–20. Philadelphia, PA: Psychology Press (Taylor & Francis).
- Burge, Tyler. 1979. “Individualism and the Mental.” *Midwest Studies in Philosophy* 4: 73–121.
- Citton, Yves. 2006. *L'Envers de la Liberté. L'invention d'un imaginaire spinoziste dans la France des Lumières*. Paris: Éditions Amsterdam.
- Dennett, Daniel. 1991. *Consciousness Explained*. Harmondsworth: Penguin Books.
- Dewey, John. 1963. *Experience and Education (1938)*. Reprint. New York: Macmillan; Collier Books.
- Dretske, Fred. 1995. *Naturalizing the Mind*. Cambridge, Mass.: MIT Press.
- Duchesneau, François. 2017. “Degrees and forms of sensibility in Haller’s physiology.” In *Human and Animal Cognition in Early Modern Philosophy and Medicine*, edited by S. Buchenau and R. Lo Presti, 201–20. Pittsburgh: University of Pittsburgh Press.

- Furth, M. 1988. *Substance, Form and Psyche: An Aristotelean Metaphysics*. Cambridge: Cambridge University Press.
- Hill, Emta. 1968. "Materialism and Monsters in the *Rêve de D'Alembert*." *Diderot Studies* 10: 67–93.
- Hull, David. 1992. "Individual." In *Keywords in Evolutionary Biology*, edited by E. Fox Keller et al., 180–7. Cambridge, Mass.: Harvard University Press.
- Husserl, Edmund. 1989. *Ideas Pertaining to a Pure Phenomenology and a Phenomenological Philosophy*. Vol. 2. Translated by R. Rojcewicz and A. Schuwer. Dordrecht: Kluwer.
- Jeannerod, Marc. 1983. *Le cerveau-machine*. Paris: Fayard.
- Kosman, Aryeh. 1987. "Animals and other beings in Aristotle." In *Philosophical problems in Aristotle's Biology*, edited by J. Lennox and A. Gotthelf, 360–91. Cambridge: Cambridge University Press.
- Lenz, Martin. 2013. "Locke as a social externalist." In *Continuity and Innovation in Medieval and Modern Philosophy: Knowledge, Mind, and Language*, edited by J. Marenbon, 53–67. Oxford: Oxford University Press.
- Menn, Stephen. 1998. *Descartes and Augustine*. Cambridge: Cambridge University Press.
- Mijuskovic, Ben Lazare. 1974. *The Achilles of Rationalist Arguments: The Simplicity, Unity and Identity of Thought and Soul from the Cambridge Platonists to Kant. A Study in the History of an Argument*. The Hague: Martinus Nijhoff.
- Morfino, Vittorio. 2006. "Spinoza: An Ontology of Relation?" *Graduate Faculty Philosophy Journal* 27(1): 103–27.
- Olson, Eric. 1997. *The Human Animal: Personal Identity Without Psychology*. Oxford: Oxford University Press.
- Pradeu, Thomas. 2012. *The Limits of the Self: Immunology and Biological Identity*. Translated by E. Vitanza. Oxford: Oxford University Press.
- Renz, Ursula. 2017. "Finite Subjects in the Ethics: Spinoza on Indexical Knowledge, the First Person and the Individuality of Human Minds." In *Oxford Handbook of Spinoza*, edited by Michael Della Rocca, 204–19. Oxford: Oxford University Press.
- Richards, Richard A. 2010. *The Species Problem: A Philosophical Analysis*. Cambridge: Cambridge University Press.
- Ricoeur, Paul. 1992. *Oneself as Another*. Translated by Kathleen Blamey. Chicago: University of Chicago Press.
- Schürmann, Reiner. 2003. *Broken Hegemonies*. Translated by Reginald Lilly. Bloomington and Indianapolis: Indiana University Press.
- Stein, Gertrude. 1995. *The Geographical History of America or the Relation of Human Nature to the Human Mind* (1936). Baltimore: Johns Hopkins University Press.
- Stevens, Wallace. 1997. *Collected Poetry and Prose*. Edited by Frank Kermode and Joan Richardson. New York: Library of America.
- Strawson, P. F. 1959. *Individuals*. London: Methuen.
- Sturm, Thomas. 2007. "The self between philosophy and psychology: The case of self-deception." In *Psychology's Territories: Historical and Contemporary Perspectives from Different Disciplines*, edited by M. Ash and T. Sturm, 169–92. Mahwah, NJ: Erlbaum.

- Thiel, Udo. 1998. "Locke and Eighteenth-Century Materialist Conceptions of Personal Identity." *Locke Newsletter* 29: 59–83.
- Thiel, Udo. 2006. "Self-Consciousness and Personal Identity." In *The Cambridge History of Eighteenth-Century Philosophy*, edited by K. Haakonssen, vol. 1, 286–318. Cambridge: Cambridge University Press.
- Thiel, Udo. 2011. *The Early Modern Subject. Self-Consciousness and Personal Identity from Descartes to Hume*. Cambridge: Cambridge University Press.
- Thiel, Udo. 2015. "Materialismus und Subjektivität bei Diderot." In *Denis Diderot. Aufklärer-Schriftsteller-Philosoph*, edited by T. Knapp and C. Pieberl, 75–94. Vienna: Löcker Verlag.
- Toto, Francesco. 2015. *L'individualità dei corpi. Percorsi nell'Etica di Spinoza*. Milan: Mimesis.
- Vygotsky, Lev. 1972. "Spinoza's Theory of the Emotions in Light of Contemporary Psychoneurology," translated by E. Berg. *Soviet Studies in Philosophy* 10: 362–82.
- Wolfe, Charles T. 2007. "Determinism/Spinozism in the Radical Enlightenment: the cases of Anthony Collins and Denis Diderot." In *Boundaries in the Eighteenth Century*, special issue of *International Review of Eighteenth-Century Studies (IRECS)*, edited by Pasi Ihalainen et al., vol. 1: 37–51.
- Wolfe, Charles T. 2010. "From Spinoza to the socialist cortex: Steps toward the social brain." In *Cognitive Architecture. From Bio-Politics to Noo-Politics*, edited by D. Hauptmann and W. Neidich, 184–206. Rotterdam: 010 Publishers, Delft School of Design Series.
- Wolfe, Charles T. 2014. "Sensibility as vital force or as property of matter in mid-eighteenth-century debates." In *The Discourse of Sensibility. The Knowing Body in the Enlightenment*, edited by Henry Martyn Lloyd, 147–70. Dordrecht: Springer.
- Wolfe, Charles T. 2016. *Materialism. A Historico-Philosophical Introduction*. Dordrecht: Springer.
- Wolfe, Charles T. 2017. "Varieties of vital materialism." In *The New Politics of Materialism. History, Philosophy, Science*, edited by S. Ellenzweig and J. Zammito, 44–65. London: Routledge.
- Yaffe, Gideon. 2004. *Manifest Activity: Thomas Reid's Theory of Action*. Oxford: Oxford University Press.

# Apperception and Affinity: Kant on the Material Condition of the Identity of the “Psychological Person”

## Introduction

In my paper I will focus on the link between Kant’s theory of personal identity and the principle of transcendental affinity. This might sound strange because, at least at first glance, the two notions seem to refer to entirely different issues.

As we will see, the doctrine of the transcendental affinity introduced in the A-Deduction of *Critique of Pure Reason* (Kant 1998) concerns the question of the genesis of sense-formation: by elaborating the affinity thesis Kant tries to grasp the lowest level in the process of combining the sensory manifold into a unified representation. These investigations have seemingly little to do with the issue of personal identity, which is usually examined in connection with Kant’s doctrine of inner sense and in the context of his polemics against rational psychology as set out in the two versions of the paralogism chapter.<sup>1</sup>

However, it is not difficult to find the nexus between the two topics if we know that, on the one hand, the affinity of the manifold is the most fundamental condition of the possibility for realising the original unity of the apperception, and, on the other, the original unity of the apperception is the most fundamental condition of the possibility for regarding ourselves as identical persons (persons in the “psychological sense” – I will explain below how this restriction of the concept of person is to be understood). Thus, the connecting link is the unity of apperception. This will be the focus of my paper; and my main question to be answered is: How must the mind articulate the outside world if there is to

---

<sup>1</sup> Out of a vast amount of material on this subject, some essential books and studies are Paton 1936, 387–425; Heimsoeth 1956, 227–57; Powell 1990; Horstmann 1993; Brook 1994, 179–207; Strawson 1997, 252–67; Ameriks 2000, 128–76; Rosefeldt 2000, 96–135; Allison 2004, 333–56; Sellars 2007, 411–36; Kitcher 2011, 180–200; Zobrist 2011; Longuenesse 2017, 140–69.

be consciousness of personal identity at all? Or to put it another way: How must the appearances “offer themselves to us” if there is to be consciousness of personal identity at all?

I will proceed as follows. I will start with a prefatory explanation of Kant’s distinction between what we might call a stronger and a weaker notion of person. I will then sketch the process of sense-formation on the basis of the A-Deduction in order to set the scene for taking a closer look at the tenet of transcendental affinity. I will then clear up the notion’s relation (A) to the reproducibility and associability of appearances; (B) to the transcendental object as a correlate of the unity of apperception; (C) to the transcendental function of the imagination; and (D) to the regulative principles of systematicity as developed in the “Appendix” to the “Dialectic”. After clarifying the tie between personal identity and the uniformity of nature, I will turn to *Critique of the Power of Judgment* (Kant 2000) in order to describe how Kant tries to reconceptualise the affinity thesis with the help of the principle of the purposiveness of nature. Finally, I will discuss some paragraphs from the so called *Opus Postumum* (Kant 1993), pointing out that a great part of the handwritten remains of the late Kant can be viewed as a struggle to reformulate the problem in a “post-critical” framework.<sup>2</sup>

## Affinity and Personal Identity in the *Critique of Pure Reason*

### Psychological and moral personality

First, I should briefly clarify an important distinction regarding the Kantian notion of person in order to define the limits of my investigations. For this reason I will take a cursory glance at the “Third Paralogism” (the paralogism of personality) in the first *Critique*. I will not analyse this passage closely, nor will I highlight the historical background from which Kant’s considerations had emerged; I will just summarise the Kantian line of thought and its results.

In the “Third Paralogism” Kant argues that the fact that all of our representations are necessarily attributed to one and the same logical subject, i.e. to the “I” in “I think” – which is the conceptual expression of the transcendental unity of apperception (cf. Longuenesse 2017, 104) – does not guarantee at all that this “I” refers to one and the same numerically identical person. The representation of this “I” encompasses no manifold within itself, it is “wholly empty of content” (A355), “we can never have even the least concept” about

---

2 Let me insert here a parenthetical remark. In this paper I use such terms as sense-formation or sense-bestowal, which, or course, cannot be found in Kantian texts – they can be found in phenomenological discourses. I do not have space here to lay the methodological ground for a phenomenological reading of the transcendental deductions. I just note that I follow such interpretations.



it (A346/B404). That is why we cannot infer the objective persistence of our identical selves from the consciousness of the unity of the contents of our thoughts.

As Kant famously puts it:

The identity of the consciousness of Myself in different times is therefore only a formal condition of my thoughts and their connection, but it does not at all prove the numerical identity of my subject in which – despite the logical identity of the I – a change can go on that does not allow it to keep its identity. (A363)

On the basis of this reasoning we have to distinguish two notions of person (and consequently two ways of thinking about personal identity). The “weaker” notion denotes nothing more than the logical unity of self-consciousness. But Kant also allows a “stronger” notion to be applied, which refers to the moral aspect of the term, and which seems to fall under the otherwise rejected rationalist concept of person. This concept

can remain (insofar as it is merely transcendental, i.e. a unity of the subject which is otherwise unknown to us, but in whose determinations there is a thoroughgoing connection of apperception), and to this extent this concept is also necessary and sufficient for practical use [...]. (A366–67)

The most concise formulation of this distinction is provided by *The Metaphysics of Morals*:

*Moral personality* is [...] nothing other than the freedom of a rational being under moral laws (whereas *psychological personality* is merely the capacity to become conscious of one’s identity in the different states of one’s existence). (Kant 1991, 50 [AA VI 223]; my emphasis)

In this paper I do not consider the question of moral personality; I only discuss the psychological aspect of the concept of person (and personality, personal identity, consciousness of personal identity). So the consciousness of personal identity (in the psychological sense) means here nothing more than the persistent unity of the consciousness in the synthesis of the manifold of representations.

## The constitution of experience

After this preliminary note I will outline the constitution of experience as delineated in the first *Critique*.<sup>3</sup> One of Kant’s main purposes in the first-edition Deduction chapter is to explore the different levels in the process of sense-formation. In the second part of the

---

3 When discussing the process of sense-formation, I consider the two versions of the Deduction chapter as complementary. I do not have space here to offer an exhaustive justification of this reading; I just refer to Longuenesse’s and Allison’s interpretations (Longuenesse 1998; Allison 2015).



third section of the A-Deduction (A119–30) he opens a bottom-up perspective and construes the way of sense-bestowal starting with the reception of the raw material of sensible impressions and ending with the moment of the conceptual determination of a unified representation.

According to the theory of the three-fold synthesis (the synthesis of apprehension, reproduction and recognition) this process can be portrayed as follows. The imagination runs through the intuitive manifold received by the sensibility, looking for connections in this unordered plethora of sensory data and trying to mobilise the appropriate schema as the rule of the synthesis of apprehension. It has to be emphasised that the schema is not an image (neither a blurred image nor a “silhouette”).<sup>4</sup> It is an “invisible form”<sup>5</sup> (it cannot be represented), a dynamic rule for mediating between the sensibility and the understanding. It is the schematising imagination that prepares and completes the sensible impressions unifying them into *one* representation (cf. Lohmar 1998, 73).

Since the impressions (as modifications of inner sense) are subject to the conditions of time (as the form of inner sense), all appearances are given successively and must be held together, that is, the preceding item has to be related to the following one in order for one to be able to comprehend the whole representation. This “holding together” is the accomplishment of the synthesis of reproduction, which is inseparable from the apprehension (A102).

The empirical synthesis has, of course, a transcendental precondition, namely the transcendental synthesis of the imagination, which can be characterised as an *a priori* unifying function standing in an indissoluble connection with the transcendental apperception. According to this correlation established by the categories, the transcendental synthesis of the imagination presupposes the original unity of apperception. However, the unity of apperception is realised only by the transcendental synthesis of the imagination (Allison 2015, 134).

Our perception stands from the very first moment under the guidance of the categories. We always perceive “extensive and intensive magnitudes” (e.g. the roundish shape and red colour of an apple) and it is only possible because the mathematical categories (the categories of quantity and quality) are from the outset involved in the process of sense-bestowal. The application of the dynamical categories (or at least the categories of relation) is also necessary in a state of “normal” perceiving, because they bind together the otherwise dispersed and separated empirical consciousnesses into a single transcendental consciousness.<sup>6</sup> And

---

4 As, for instance, Maurizio Ferraris falsely maintains: Ferraris 2013, 82.

5 I borrowed this expression (applied in connection with the Kantian schemata) from Tamás Ullmann: Ullmann 2010, 118.

6 Cf. Wenzel 2005; Allison 2015, 420–6. One might argue against these statements by referring to certain passages in §13 of the *CPR* where Kant writes that “objects can indeed appear to us without necessarily having to be related to functions of the understanding, and therefore without the understanding containing their a priori conditions”; and again: “Appearances would nonetheless offer objects to our intuition, for intuition by no means requires the functions of thinking.” (A89–91/B122–3). However, these references do not contradict what I maintained. There is no need for application of the

since Kant identifies the categories as the logical functions of judgement (B143), it is not difficult to see that the whole process has a strictly teleological structure, the final stage of which is the synthesis of recognition, the act of subsuming a sensible representation under an empirical concept in an objectively valid judgement.

## Transcendental affinity

One of the most exciting questions of the A-Deduction is the following: what guarantees that our imagination is able to discover any kind of regularities in the plenitude of sense-data we receive? If there were no explorable connections between the items of a chaotic aggregate of sensible impressions, it would be impossible to generate schemata at all. We would have no rules for the apprehension of the manifold, consequently no concepts, no unity of apperception and, of course, no consciousness of personal identity. The raw material of the sensory data would not be in conformity with transcendental apperception. I quote a famous passage from the A-Deduction:

If cinnabar were now red, now black, now light, now heavy, if a human being were now changed into this animal shape, now into that one, if on the longest day the land were covered now with fruits, now with ice and snow, then my empirical imagination would never even get the opportunity to think of heavy cinnabar on the occasion of the representation of the color red. (A100)

There must thus be a rule, Kant argues, “to which the appearances are already subjected in themselves” (A101). Of course, it often happens that the sensible impressions pose a real challenge to our imagination; for example, bad sight conditions might open up a broad field for associations and various competing schemata. But we never call into question (in normal waking states, not in altered states of consciousness) that the manifold we receive from the outside world can be structured and unified; that we do have or we are able to generate the proper procedural rules according to which the appearances could be apprehended.

“There must therefore be an objective ground,” due to which it is “impossible for appearances to be apprehended [...] otherwise than under the condition of a possible synthetic unity of this apprehension” (A121); and this objective ground is the “transcendental affin-

---

categories in the case of one’s unconscious representations (like Leibniz’s micro-perceptions), nor in the case of the animal “perception” (*apprehensio bruta* – cf. Kant’s letter to Beloselsky, Summer 1792, AA XI 344–6), not to mention God’s intellectual intuition (B139, 145). But I do not have to discuss these topics because the issue at stake here is human perception (*Wahrnehmung*), which is a conscious act involving the categories. For an interpretation of Kant’s views on animal sensation, see Naragon 1990. For interpretations of Kant’s views on intellectual intuition (or intuitive understanding), see Förster 2012, 143–52; Leech 2014.

ity, of which the empirical affinity is the mere consequence” (A113–14). How is this to be conceived?

In order to answer this question, I need to clarify the relation of affinity to several other concepts, namely (A) to the reproducibility and associability of appearances; (B) to the transcendental object as the correlate of the unity of apperception; (C) to the transcendental function of the imagination; and (D) to the regulative principles of systematicity as discussed in the “Appendix” to the “Dialectic”.

(A) Kant defines affinity as “the ground of the possibility of the association of the manifold” and claims that all appearances “are associable in themselves and subject to universal laws of a thoroughgoing connection in reproduction” (A113, A122). An appearance is associable if and only if its manifold is taken as belonging together in a determinate way: it has similar phenomenal properties whenever it appears to us. It can be associated with other representations because our imagination “internalised” its sensible image and is able to reproduce it. In this regard, associability and reproducibility can be considered as equivalent; they both presuppose the regularity, that is to say, the affinity of appearances. The somehow obscure statement “all appearances are associable in themselves” asserts that the sense-data can only be apprehended as combined into potentially identifiable perceptual unities.

(B) This statement (“all appearances are associable in themselves”) also implies the seemingly trivial insight that the successively given representations of one and the same object have to be related to one and the same object; otherwise the sensible impressions could not be bound together in the apprehension. But this one and the same object is itself not a representation. It is a mere correlate of the unity of apperception: the inaccessible, non-empirical ground for the unity in the manifold of cognition. Kant calls this the transcendental object. It does not “contain any determinate intuition at all”, and is “only the representation of appearances under the concept of an object in general” (A109, A251). We are now in a position to understand the link between transcendental affinity, the transcendental object and the transcendental unity of apperception. If there were no regularity (no affinity) in the sensory manifold, we could not relate the representations of a “something” to the same transcendental object, that is to say, the basic criterion for the formal unity of consciousness would fall apart.

(C) According to the Kantian doctrine, the ultimate source of affinity is nothing other than the transcendental function of the imagination, which “installs” the regularity into our impressions so that they can be apprehended in agreement with and under the governance of the transcendental structure of our consciousness. As Kant puts it:

It is therefore certainly strange, yet from what has been said thus far obvious, that it is only by means of this transcendental function of the imagination that even the affinity of appearances, and with it the association and through the latter finally reproduction in accordance with laws, and consequently experience itself, become possible; for without them no concepts of objects at all would converge into an experience. (A123)

This is a deeply problematic contention. It is fiercely disputed among Kant scholars whether transcendental idealism can offer an explanation of the thoroughgoing affinity of appearances (Allison 1974; Beck 1981; Westphal 1997; Westphal 2004, 68–126; Szegedi 2007, 98–107; Papp 2010, 99–104; Ullmann 2010, 47–55; Allison 2015, 264–71). If things in themselves are already not structured to a certain degree, it is hard to see how our transcendental imagination could adjust the matter of sensations to the demands of their apprehensibility without violation of the autonomy of our sensibility. But if the transcendental affinity is to be attributed to the matter of sensations (and so eventually to the objects in themselves), then the synthesis functions of the understanding have nothing more to do than to reconstruct, not construct the order of appearances (Westphal 2004, 90). In this case our understanding, contrary to what Kant maintains, is not “itself the legislation for nature” (A126) because there is a non-subjective, formal and material condition for the possibility of experience, which is an untenable assertion in the critical framework (Westphal 2004, 125).

Both interpretations may jeopardise the boundaries of transcendental idealism. It is highly remarkable that Kant erases all traces of the affinity doctrine from the second-edition Deduction chapter, maybe because he realises this jeopardy (Beck 1981, 457). Anyway, Kant’s tacit assumption must be that because of the noumenal conditions of all phenomena nothing can affect our capacity for representation in a completely chaotic way. The data of senses we receive must be pre-structured in some measure so as to meet the requirements of being apprehensible, reproducible and associable.

(D) I also have to clear up the tie between the affinity concept elaborated in the A-Deduction, on the one hand, and the affinity concept introduced in the “Appendix” to the “Dialectic” (in connection with the regulative principles of the systematic unity of experiences), on the other. As we have seen before, the former one, the affinity of appearances is an accomplishment of the transcendental function of the imagination and consequently is within the purview of the understanding. The latter notion of affinity denotes the law of the kinship of all concepts and is a principle of reason (not of understanding); a principle “which offers a continuous transition from every species to every other through a graduated increase of varieties” (A658/B686).

At the first glance these two concepts of affinity have little to do with each other. But we have to bear in mind the broader context, namely that the transcendental affinity of the sensible manifold manifests itself in the fact that we are able to apply schemata as rules for apprehending the impressions we perceive. But schemata are results of logical operations: we compare representations with one another; we reflect the similarities among the elements and abstract from the dissimilarities (cf. *Logik*, AA IX 94–5). So it is hard to escape the conclusion that the two affinity concepts mutually imply each other (to a certain extent): the affinity of appearances entails the comparability, classifiability and conceptual determinability of our representations (and vice versa).

Now, the classifiability of the representations is just a presupposition, albeit a subjectively necessary transcendental presupposition. As Kant famously puts it:

If among the appearances offering themselves to us there were such a great variety [...] that even the most acute human understanding, through comparison of one with another, could not detect the least similarity (a case which can at least be thought), then the logical law of genera would not obtain at all, no concept of a genus, nor any other universal concept, indeed no understanding at all would obtain, since it is the understanding that has to do with such concepts. The logical principle of genera therefore presupposes a transcendental one if it is to be applied to nature [...]. According to that principle, sameness of kind is necessarily presupposed in the manifold of a possible experience (even though we cannot determine its degree a priori), because without it no empirical concepts and hence no experience would be possible. (A654/B682)

Kant argues as follows. The whole of nature constitutes a system inasmuch as it is subject to the transcendental principles and laws of sensibility and understanding. For this very reason experience must also constitute a system of possible empirical cognitions (because “all empirical laws are only particular determinations of the pure laws of the understanding” [A128]). So if there is to be any experience at all, then all appearances must stand in thoroughgoing affinity, on the one hand; and our empirical concepts must fit into a logical system, on the other. And the connection between these two kinds of affinity might be described like this: the systematic unity of experience is the affinity of appearances considered as brought under the categories and organised by the reason. (cf. Allison 2015, 275: 60n)

After having briefly discussed the elements of this conceptual net around affinity, we now have to pose the question: What follows from all this regarding Kant’s view on personal identity? We have seen that affinity is the most basic condition for realising the transcendental unity of the apperception and consequently the consciousness of the identity of a person’s self. In the light of the analysis presented above one might contend that if there is to be consciousness of personal identity at all, all of our sensations we receive from the outside world must be suitable for being articulated as *possible* experiences potentially fitting into a systematic (though undetermined) unity of nature.

This assertion does not mean that whatever appears to us must be brought under empirical concepts, nor does it mean that we cannot err regarding what we have perceived. It only means that the sensory manifold must be suitable for being apprehended as extensive and intensive magnitudes in the formal unity of consciousness. To put it the other way around: if our sensible impressions cannot be articulated as possible experiences, this means that the formal unity of consciousness cannot be realised in the synthesis of the manifold of representations, that is to say, there is no “standing or abiding self” in the stream of the modifications of the mind (A107).

# Affinity and Personal Identity in the *Critique of the Power of Judgment*

I have already mentioned that in the second-edition Deduction chapter Kant completely ignores the possible hazard of a “transcendental chaos”, probably because of the ambiguity of the doctrine of affinity. But the problem did not let him rest. He reconceptualised it as the subjectively necessary principle of the purposiveness of nature in *Critique of the Power of Judgment*. He also reconsidered the question of which cognitive faculty is responsible for establishing the affinity. In the first *Critique*, as we have seen, the affinity of the appearances was generated by the transcendental function of the imagination, while the regulative principle of the affinity of all concepts was presupposed by the reason, which had to organise the cognitions of the understanding into a systematic unity of knowledge. Both aspects of the affinity can be discerned in the third *Critique*, too; but this time the principle is under the purview of the reflecting power of judgment. I quote from the first introduction:

In our power of judgment we perceive purposiveness insofar as it merely reflects upon a given object, whether in order to bring the empirical intuition of that object under some concept (it is indeterminate which), or in order to bring the laws which the concept of experience itself contains under common principles. Thus the power of judgment is properly technical; nature is represented technically only insofar as it conforms to that procedure of the power of judgment and makes it necessary. (Kant 2000, 22 [AA XX 220])

Kant speaks here about two kinds of *formal or logical* purposiveness – neither of them has to do anything with *objective* purposiveness, i.e. with natural ends (cf. Wachter 2012, 52).

(1) He speaks about the logical operation of reflecting upon an object in order to subsume the empirical intuition under some indeterminate concept, and he explains further that we perceive purposiveness insofar as the disposition of the imagination and the understanding is well-proportioned with respect to cognition in general (Kant 2000, 123 [AA V 238–39]). This means nothing more than that these two faculties start to mutually animate each other, while realising the formal unity of the consciousness in the synthesis of the manifold of representations. I think this kind of purposiveness precisely corresponds to the affinity of appearances as characterised in the A-Deduction.

(2) The second kind of formal purposiveness mentioned in the above quotation is the reformulation of the principle of the affinity of all concepts. Reflecting upon an object “in order to bring the laws which the concept of experience itself contains under common principles” is exactly the same operation which is already familiar from the “Appendix” to the “Dialectic” chapter: it is the integration of our cognition into a systematic unity of knowledge.

So within the conceptual framework of *Critique of the Power of Judgment* we might conclude that if there is to be consciousness of personal identity at all, the sensible manifold must be suitable for bringing about such a proportion of the disposition of imagination and understanding in which these two cognitive powers can mutually enliven each other so as to constitute a possible experience.

As we can see, in the third *Critique* Kant begged the question about the specific characteristics of the contents of appearances, namely that they must display some minimal degree of recognisable regularity without which there is no experience at all; and his focus shifted to the relationship or disposition of the cognitive faculties. However, Kant never ceased to ruminate on the question: what is the ultimate guarantee for the conformity of the content of our sensations with the transcendental structure of our consciousness.

## Affinity and Personal Identity in the *Opus postumum*

I think a great part of *Opus postumum*, especially the sections of the so-called “Ether Deduction” can be regarded as a repeated attempt to elaborate the answer to the aforementioned question. In these passages Kant obviously goes beyond the boundaries of his critical system insofar as he overtly introduces a material transcendental condition of the possibility for experience.

I cannot get into details about the whole transition project, which aims to bridge the gap between the metaphysical foundations of natural science and physics (cf. Friedman 1992, 290–342; Rollmann and Hahmann 2011; Hall 2015). I only concentrate on some paragraphs from the second, fifth and twelfth fascicles in order to reconstruct the main argumentative steps of the “Ether Deduction” and to point out its connection with the question of affinity.

Kant’s line of thought can be summarised as follows:

(1) If there is to be experience at all, there must be only one experience, “in which all perceptions are represented as in thoroughgoing and lawlike connection” (we already know it from the first *Critique*: A110). When speaking about experiences, we understand thereby only “a continuous series of possible perceptions,” that is, a series of conscious (and non-conscious) subjectively valid representations, which can potentially be raised to the objective validity of a single all-encompassing experience. There can be no gap between our representations, otherwise this gulf (*Kluft*) would prevent the transition from one appearance to another. They could not be combined with each other, hence “the unity of the guiding thread of experience would be torn apart” and the transcendental unity of the apperception would be dissolved (Kant 1993, 88 [AA XXII 552]).

(2) If there is only one experience, there can be no empty space at all. Kant argues as follows:



Intuitions in space and time are mere forms, and, lacking something which renders them knowable for the senses, furnish no real objects whatsoever to make possible an existence in general [...]. Consequently, space and time would be left completely empty for experience. (Kant 1993, 68 [AA XXI 217])

In Kant's view the atomistic explanation is untenable because this implies that space is "a composition of the *full* with the *void* between it" (Kant 1993, 69 [AA XXI 218]); and this entails that we should be apprised of the existence of a matter through the intermediary empty space between and around the atoms, which is self-contradictory because empty space is not an object of possible experience. "The principle for the agreement of all perceptions with the conditions of the possibility of experience excludes any void, since this is not an object of possible experience." (Kant 1993, 81 [AA XXI 552]). For this reason, space must be filled with a universally distributed, all-penetrating matter, which forms a continuum.

(3) This world-material or primordial material, call it ether or caloric, has to be an incessantly self-agitating matter attracting and repelling "in its own parts" (Kant 1993, 73 [AA XXI 225]); otherwise (without this restless vibration) it could not stimulate the sense organs, thus no perception of any object would take place (Kant 1993, 87 [AA XXII 551]). Thus, on the one hand, the ether as a substrate of all dynamic forces is the material ground for all physical bodies: the "figure and texture of all bodies" rests upon this primordial moving matter (AA XXI 548).<sup>7</sup> On the other hand, the ether is the ultimate source of perceptual affection; and as a homogeneous continuum it is the basis for the unification of all perceptions into an absolute whole of a single experience.

As you can see, Kant tries to prove his claim by demonstrating the impossibility of its contrary:

To carry out this indirect mode of proof – which is not objective, from experience (empirical), but from the principle of the possibility of experience in general (*a priori*), and consequently subjective – appears strange; for such a mode of inference does not seem at all consistent or possible. [...] [The actuality of a universally distributed, all-penetrating world-material] rests solely on the principle of the possibility of outer experience and is thus known and confirmed *a priori*, according to the principle of identity. For, without presupposing this material, I could not have any outer experience at all: Empty space is not an object of possible experience (Kant 1993, 74, 76 [AA XXI 226, 229]).

The mere possibility of experience already guarantees enough; moreover, it alone guarantees the reality of this material which fills all spaces. (Kant 1993, 70 [AA XXI 220])

---

<sup>7</sup> This passage was omitted in the English translation.



I think Kant's final answer to the question of the source of affinity is that there is an all-penetrating, homogeneous primordial material, the ether, which is the structural substrate of all matter, and it constitutes the pre-structure of all appearances through its motion. The ether is the transcendental material condition due to which the sensible manifold is always apprehensible, associable, reproducible, conceptually determinable and classifiable. The ether is the most fundamental condition of the possibility for the original unity of the apperception and consequently for the consciousness of personal identity, too.

## Conclusion

Let me end with a few concluding remarks. In this paper I have tried to provide a comprehensive overview of Kant's endeavours to answer the question: How must the appearances "offer themselves to us" if there is to be something like an identical "psychological personality"? I have found three answers from Kant's critical and "post-critical" period.

In the first *Critique* Kant argued for the thoroughgoing transcendental affinity of appearances. This account is utterly ambiguous inasmuch as it can be laid out in two ways. On the one hand, if the transcendental imagination alone takes the full burden for injecting the affinity into the sensory data, then the imagination must violate the autonomy of the sensibility. On the other hand, if the affinity lies in the object, then there is a formal and material condition of the possibility for experience, which is a statement that Kant certainly would have rejected at the time of the first *Critique*. This indissoluble ambiguity and the theological connotation of the concept might be the main reasons why Kant erased the notion of affinity from the second-edition Deduction chapter.

In *Critique of the Power of Judgment* Kant rephrased the problem as the subjectively necessary transcendental principle of nature's purposiveness for our faculty of cognition. Kant's implicit assertion was that the sensible manifold is always suitable for bringing about such a proportion of the disposition of our imagination and understanding in which the two cognitive powers can mutually animate each other. Kant begged the question here concerning what the basis for this suitability of the manifold is.

I tried to point out that Kant also outlined a third approach in his *Opus postumum*. He did not shrink from reconceptualising his former thoughts in a way which actually bursts the framework of his transcendental idealism insofar as he argued for a transcendental material condition of the possibility for experience. On the basis of these handwritten remains we can decipher his late view on the question of the basic condition of the possibility for realising a psychological personality, too. One might put it this way: if there is to be consciousness of personal identity at all, the ether must exist.

# Bibliography

## Primary Sources

- Kant, Immanuel. 1900ff. *Gesammelte Schriften*. Vols 1–22, edited by Preussische Akademie der Wissenschaften. Berlin: George Reimer, later Walter de Gruyter; vol. 23., edited by Deutsche Akademie der Wissenschaften zu Berlin. Berlin: Walter de Gruyter; from vol. 24., edited by Akademie der Wissenschaften zu Göttingen. Berlin: Walter de Gruyter [AA, volume, page]
- Kant, Immanuel. 1998. *Critique of Pure Reason*. Edited and translated by Paul Guyer and Allen W. Wood. New York: Cambridge University Press.
- Kant, Immanuel. 2000. *Critique of the Power of Judgment*, edited by Paul Guyer. Translated by Paul Guyer and Eric Matthews. New York: Cambridge University Press.
- Kant, Immanuel. 1991. *The Metaphysics of Morals*. Translated by Mary Gregor. New York: Cambridge University Press.
- Kant, Immanuel. 1993. *Opus postumum*. Edited by Eckart Förster, translated by Eckart Förster and Michael Rosen. New York: Cambridge University Press.

## Secondary Sources

- Allison, Henry E. 1974. “Transcendental Affinity: Kant’s Answer to Hume.” In *Kant’s Theory of Knowledge: Selected Papers from the Third International Kant Congress*, edited by Lewis White Beck, 119–27. Dordrecht: D. Reidel Publishing Company.
- Allison, Henry E. 2004. *Kant’s Transcendental Idealism: An Interpretation and Defense*. Revised and enlarged edition. New Haven: Yale University Press.
- Ameriks, Karl. 2000. *Kant’s Theory of Mind: An Analysis of the Paralogisms of Pure Reason*. New York: Oxford University Press.
- Beck, Lewis W. 1981. “Kant on the Uniformity of Nature.” *Synthese* 47: 449–64.
- Brook, Andrew. 1994. *Kant and the Mind*. New York: Cambridge University Press.
- Ferraris, Maurizio. 2013. *Goodbye, Kant! What Still Stands of the Critique of Pure Reason*. Translated by Richard Davies. Albany: State University of New York Press.
- Förster, Eckart. 2012. *The Twenty-Five Years of Philosophy: A Systematic Reconstruction*. Translated by Brady Bowman. Cambridge: Harvard University Press.
- Friedman, Michael. 1992. *Kant and the Exact Sciences*. Cambridge: Harvard University Press.
- Hall, Bryan W. 2015. *The Post-Critical Kant: Understanding the Critical Philosophy through the Opus postumum*. New York: Routledge.
- Heimsoeth, Heinz. 1956. “Persönlichkeitsbewußtsein und Ding an sich in der Kantischen Philosophie.” In *Studien zur Philosophie Immanuel Kants: Metaphysische Ursprünge und Ontologische Grundlagen*, edited by Ingeborg Heidemann, 227–57. Köln: Kölner Universitäts-Verlag.
- Horstmann, Rolf-Peter. 1993. “Kants Paralogismen.” *Kant-Studien* 84(1): 408–25.
- Kitcher, Patricia. 2011. *Kant’s Thinker*. New York: Oxford University Press.

- Leech, Jessica. 2014. "Making Modal Distinctions: Kant on the Possible, the Actual, and the Intuitive Understanding." *Kantian Review* 19(3): 339–65.
- Lohmar, Dieter. 1998. *Erfahrung und kategoriales Denken: Hume, Kant und Husserl über vorprädikative Erfahrung und prädikative Erkenntnis*. Dordrecht: Springer-Science+ Business Media.
- Longuenesse, Béatrice. 1998. *Kant and the Capacity to Judge. Sensibility and Discursivity in the Transcendental Analytic of the Critique of Pure Reason*. Translated by Charles T. Wolfe. Princeton: Princeton University Press.
- Longuenesse, Béatrice. 2017. *I, Me, Mine: Back to Kant, and Back Again*. New York: Oxford University Press.
- Naragon, Steve. 1990. "Kant on Descartes and the Brutes." *Kant-Studien* 81(1): 1–23.
- Papp Zoltán. 2010. *Elidőzni a szépnél: Kant esztétikájáról*. Budapest: Atlantisz.
- Paton, Herbert J. 1936. *Kant's Metaphysic of Experience: A Commentary on the First Half of the Kritik der reinen Vernunft*. Vol. 2. London: George Allen & Unwin Ltd.
- Powell, C. Thomas. 1990. *Kant's Theory of Self-Consciousness*. New York: Oxford University Press.
- Rollmann, Veit-Justus and Hahmann, Andree. 2011. "Weltstoff und absolute Beharrlichkeit: Die Erste Analogie der Erfahrung und der Entwurf Übergang 1–14 des Opus postumum." *Kant-Studien* 102(2): 168–90.
- Rosefeldt, Tobias. 2000. *Das logische Ich: Kant über den Gehalt des Begriffes von sich selbst*. Berlin: 2000 PHILO Verlagsgesellschaft.
- Sellars, Wilfrid. 2007. "«...this I or he or it (the thing) which thinks...»: Immanuel Kant, Critique of Pure Reason (A346; B404)." In *In the Space of Reasons: Selected Essays*, edited by Kevin Scharp and Robert B. Brandom, 411–36. Cambridge: Harvard University Press.
- Strawson, Peter F. 1997. *Entity and Identity and Other Essays*. New York: Oxford University Press.
- Szegedi Nóra. 2007. *A magában való dolog fenomenológiája: Egy klasszikus kanti probléma Husserl transzcendentális fenomenológiájának perspektívájából*. Budapest: ELTE BTK Filozófiai Intézet–L'Harmattan Kiadó–Magyar Filozófiai Társaság.
- Ullmann Tamás. 2010. *A láthatatlan forma. Sematizmus és intencionalitás*. Budapest: L'Harmattan.
- Wachter, Alexander. 2006. *Das Spiel in der Ästhetik. Systematische Überlegungen zu Kants Kritik der Urteilskraft*. Berlin: Walter de Gruyter.
- Wenzel, Christian H. 2005. "Spielen nach Kant die Kategorien schon bei der Wahrnehmung eine Rolle? Peter Rohs und John McDowell." *Kant-Studien* 96(4): 407–26.
- Westphal, Kenneth. 1997. "Affinity, Idealism, and Naturalism: The Stability of Cinnabar and the Possibility of Experience." *Kant-Studien* 88(2): 139–89.
- Westphal, Kenneth R. 2004. *Kant's Transcendental Proof of Realism*. New York: Cambridge University Press.
- Zobrist, Marc. 2011. *Subjekt und Subjektivität in Kants theoretischer Philosophie: Eine Untersuchung zu den transzendentalphilosophischen Problemen des Selbstbewusstseins und Daseinsbewusstseins*. Berlin: Walter de Gruyter.

# SELF-KNOWLEDGE



# Richard Burthogge's Epistemology and the Problem of Self-Knowledge

## Introduction

This paper focuses on the epistemology developed by Richard Burthogge, the lesser-known seventeenth-century English philosopher, and author, among other works, of *Organum Vetus & Novum* (Burthogge 1678) and *An Essay upon Reason and the Nature of Spirits* (Burthogge 1694). Burthogge's ideas had a minimal impact on the philosophy of his time, and have hitherto not been the subject of a detailed study – one short dissertation and only a few papers have been published about his thought.<sup>1</sup> This situation has reflected, to some extent, the long-lasting lack of interest among scholars in less important doctrines. While the most influential thinkers of the early modern age have been thoroughly studied, there has been surprisingly little research on the so-called *minorum gentium* philosophers. Instead of finding their way through a maze of secondary figures operating in a multitude of linguistic, socio-political and cultural contexts, the authors of synthetic accounts of seventeenth- and eighteenth-century philosophy usually preferred to focus their attention on the great names of the day, such as Leibniz, Spinoza, Locke or Berkeley. Consequently, the minor doctrines, by their nature less original and interesting, disappeared under the shadow of great doctrines – partly under the weight of comparative difficulties, and partly due to the natural hierarchy of importance. Thus, however, a multi-dimensional mosaic of doctrines forming the original landscape of early modern philosophy was replaced by an archipelago of separate theories – more or less closely related to each other. Accordingly, although the general topography of early modern thought is perfectly known, the created map contains only a few main points scattered at a considerable distance from each other. At the same time, we lose sight of the internal logic and dynamics of the whole system; for

---

1 For the most important of these works see: Ayers 2005) Grünbaum 1939; Landes 1920, 1921; Lyon 1888, 72–96; Sgarbi 2012; see also: Cassirer 1922, 543–53; Nuchelmans 1983, 117–19; Šoć 2012.

the first thing required to trace a line of development of any historical process is to identify and describe in detail all its relevant components, small as well as large ones – and this is precisely what we lack in the case of early modern philosophy.

For this reason, the need for a deeper analysis of the minor philosophers of the time has been recently emphasised, so that a new history of early modern philosophy could also include the less important doctrines (for example by Rutherford 2006, 1–4). Of many paths needed to be traversed to achieve this goal, that is, to supplement the existing map of “metropolises” with some “small towns” and “villages”, one of the most interesting leads through an unfrequented Burthoggean route. The reason for this becomes clear immediately if one considers the evolution line of early modern philosophy in its most general form, in which it is simplified to the axis Descartes – Kant, with Kantian transcendentalism seen as an extreme variant of the Cartesian philosophy or as Cartesianism taken to the ultimate epistemological consequences. Against this background, Burthogge’s writings contain a highly original concept of idealistic constructivism, developed in the post-Cartesian theoretical context, and anticipating (all proportions kept) Kant’s idealism (see, for example, Landes 1920, 1921). Moreover, his epistemology has also some interesting implications – once again Kantian in spirit – for the issue of self-interpretation and self-identity. Therefore, the detailed examination of the coincidence of ideas between Burthogge and Kant can not only deepen our understanding of early modern British philosophy, but may also throw new light on the inner logic of development of the whole of post-Cartesian epistemological thought.

The paper is in two parts. In the first and more extensive one, I outline the basic assumptions and principles of Burthogge’s epistemology (as presented in Burthogge 1694). In the second part I discuss the general implications of his theory for the problem of self-knowledge and personal identity.

## **Part One: Burthogge’s Epistemology**

### **Act of cognition**

At first glance, Burthogge’s epistemology may not seem to differ significantly from the traditional models of medieval and ancient philosophy. The only means the human mind has at its disposal to know the external reality are its three “cogitative” faculties or powers, that is “reasoning or intellection”, “sensation” and “imagination” (Burthogge 1694, 3–4). What is novel in Burthogge’s epistemological position is a remarkable idea of structural and functional uniformity among them – all three faculties are structurally and functionally similar or even isomorphic (Burthogge 1694, 3–4). Consequently, every intellectual and sensuous act of cognition or, the same thing put in Burthogge’s own terms, of “cogi-

tation” (in a broad sense of the term, see below) or “knowledge” contains in its structure three isomorphic elements or aspects.

The first of them is defined as “apprehension” (Burthogge 1694, 4), a term firmly rooted as far back as the scholastic idea of *apprehensio simplex* (considered the first operation of the mind) and by the well-established usage in seventeenth-century logic and philosophy associated with the concept of pure, judgment-free grasp of an object by the intellect. Under one name or another, the idea of such an act underlay the theories of the most influential philosophers of the time. It plays, for example, a prominent role in Descartes’ account of intellectual cognition (see, for example, his concept of intellectual *perceptio* in Descartes 1644, 12–13). No less importantly, given the unprecedented impact of the book, it appears also among the basic concepts of the Cartesian-inspired *Port Royal Logic*.<sup>2</sup>

Apprehension is an “act”. While Burthogge uses this expression most frequently to describe one particular kind of apprehending, namely the intellectual one (Burthogge 1694, 23–24), a structural-functional similarity of the cognitive powers allows us to extend it to all types of cognition, and consequently to cognition as such. Furthermore, it is precisely thanks to an act of apprehension that every cognition, intellectual as well as sensuous, takes the form of an act.<sup>3</sup> Apprehension is, therefore, the most basic and most central operation performed by the mind while knowing. It constitutes the structural core of cognition and determines its specific ontological form. However, if cognition is to be characterised as “apprehension”, it is not merely because of its act-structure, but because of its “reference to the *object*, which is known” (Burthogge 1694, 4). Apprehension is, therefore, by its very definition, an object-directed act. The proper object of apprehending referred to here is an external, out-of-mind thing, as is clear from the context (Burthogge 1694, 3–4). At the same time, it is only its indirect or, as Burthogge calls it, “ultimate” object (Burthogge 1694, 72–73). The reason for this is that, as we will see below, an act of apprehension is always mediated by “conceptions” (Burthogge 1694, 4), that is by certain conceptualisations in the form of images or notions, which Burthogge also considers to be the objects, namely the “immediate” ones, of apprehension.<sup>4</sup>

As each cognitive operation incorporates an object-directed apprehension, every act of cognition, intellectual as well as sensuous, can be considered an intentional act in a broad sense of the term. There are at least two immediate consequences of this claim. Firstly, every cognitive content has some objective, external reference. Secondly, neither mind nor object can be reduced to a pure stream of sense data, but on the contrary they are to be considered ontologically independent from the content of a cognitive act. These conclusions may, in turn, clarify the meaning of the term “perception”, used in a scanty definition of

---

2 Arnault and Nicole had used the French word “concevoir” in their original work (Arnault and Nicole, 1662, 27). The term was, however, usually translated as “apprehensio” or “apprehension” in Latin and English editions of the book, respectively; see, for example, Arnault and Nicole, 1674, 1; 1685, 41–2.

3 Here and elsewhere the term “sensation” includes “imagination”, the latter, as we will see below, being a special kind of the former.

4 See, for example, the title of the third chapter of *Essay*: “Of notion, the *immediate object of apprehension*” (Burthogge 1694, 51).



apprehension provided by Burthogge at the beginning of his *Essay* (“conscious perception”, Burthogge 1694, 4). It must be interpreted as a dynamic act of the mind, akin to “seeing” (Burthogge 1678, sect. 5; 1694, 23), and not as passive, inert reception, thereby resembling Descartes’ *perceptio* rather than Locke’s perception.

Two features make Burthogge’s approach to apprehension unique compared with those of the most prominent philosophers of his time. The first is a deliberate and systematic extension of this concept to cover sensation and imagination, and not only intellectual cognition, as for example with Descartes, thereby giving act-structure, object-direction and intentionality to all kinds of cognition. At the same time, once again contrary to the Cartesian views, apprehension of an object, that is a pure grasp of it, represents only one of the aspects of Burthogge’s cognition and cannot be performed independently as a distinct cognitive operation.

Thus every cognition is not only apprehension, but also “cogitation” (a term undoubtedly alluding to Cartesian philosophy). The word is used by Burthogge in two senses. In its narrow, technical sense it is defined as “conscious affection”, that is “affection with consciousness of that affection” (Burthogge 1694, 4). Each cognition is, therefore, a conscious act – with “consciousness” understood here as the most elementary awareness present in every conscious, even the most rudimentary, mental state (Burthogge 1694, 4–8). Owing to the semantic connection between “cogitation” and consciousness, the term can also be, and often is, used more broadly to denote a cognitive act as a whole (see, for example, Burthogge 1694, 58–60).

All that has been said so far may still look like a traditional, even Scholastic, account of cognition. What is truly revolutionary about Burthogge’s epistemology – or rather what could have been truly revolutionary, if it had been noticed by his fellow philosophers – is the idea that an external object of cognition is never presented to the mind directly, as it is in itself, but under some special form or manner of conceiving – “*modus concipiendi*” – specific to the human cognitive faculties due to their immanent structure (Burthogge 1694, 56).

An important preliminary remark must be made before discussing this part of Burthogge’s philosophy in detail. Although being an essential component of each cognitive act, the perception as described above, that is as an act of the mind directed towards an external thing, does not constitute an initial stage of cognition. In actual fact, it is only a secondary phenomenon. For Burthogge as for Locke, by whom he was clearly influenced in this respect, it is the “affection” of the sense organs, “caused” or “occasioned” by an external thing, that triggers the whole machinery of cognition:<sup>5</sup>

In all *acts of sensation* there is first an *affection* of the organ, and then a *perception* of that affection by the soul; or rather, a perception excited in the *soul* by means of

---

5 This is also true of intellect, which is not an independent source of knowledge, parallel to the sense. The former works only with the material obtained from the latter: “the understanding converses not with things ordinarily but by the intervention of the sense” (Burthogge 1694, 60). Burthogge’s theory of cognition, much as that of Kant, is therefore a hierarchical or at least sequential one.

that affection [...] a soul cannot but by *means* of *organs*, take any notice of external objects, nor the organ be a means of conveying any notice to the soul, but by being first *affected* itself. (Burthogge 1694, 152)

Accordingly, the most basic definition of “knowledge” or “cogitation” is the “conscious affection” (Burthogge 1694, 4). Nevertheless, affection (or synonymously: “impression”, Burthogge 1694, 5–8) is not to be understood as a cognitive content of a kind – some raw, external *datum* – provided for the mind through sense organs and then cognitively processed or simply absorbed into knowledge. Not only is the affection not a component of knowledge, it is not as such present in the mind. In fact, the term denotes merely the cognitive stimulus provided by the external object with the aim of activating the mind. It refers to the stimulation of the mind by the external world, or, which is the same thing viewed from the other side, to the mind’s being stimulated (affected) in some way. In this sense and only in this sense, the external thing can be said to “excite” something in the mind (Burthogge 1694, 70).

Having been affected, the mind responds actively to “form” (Burthogge 1694, 52, 56) or “frame” (Burthogge 1694, 64, 93) a “conception” – when regarded as a structural component of knowledge – or a “modification” – when considered from the ontological point of view (Burthogge 1694, 6–7). Thus we reach the third of the aforementioned features of the human cognitive faculties as characterised by Burthogge, namely their “conceptivity” (Burthogge 1694, 3). The human mind is essentially active and creative in cognition. Far from being a passive absorbent of the external data, it reacts to the stimulus provided from outside with an act of creative conceiving or, even more precisely, conceptualising. This is the proper meaning of Burthogge’s term “conception”, referring both to the process, that is to an “act of conception” (Burthogge 1694, 5), and to the product of such conceptualisation. Conception, so understood, is a spontaneous and automatic process. For the mind, forming conceptions is simply a way of acting – a manifestation of its internal mechanics. In other words, it is inherent in the cognitive faculties that they cognise by conceiving (Burthogge 1694, 3–4, 56). For structural reasons conception is also prior to any conscious act (Burthogge 1694, 71).<sup>6</sup>

As implied by the idea of structural and functional uniformity among the cognitive faculties, the conception is a universal mechanism of cognition:

It is as proper to say, that the sense and imagination do conceive, as that the reason or understanding doth; the former does as much conceive images and sentiments, as the latter does ideas and notions. (Burthogge 1694, 4)

Although essentially identical in their conceptive way of functioning, the cognitive faculties differ from one another in their particular manner of conceiving; otherwise they could

---

<sup>6</sup> According to Burthogge’s view, it is only a diversity of conceptions that enables consciousness to arise (Burthogge 1694, 4–6).

not be considered different faculties at all. The difference concerns the formal means used by them in the process of conceptualisation. While reason conceives by means of “ideas or notions”, sense and imagination do the same by means of “images”:

*Sense*, (by which I mean the power of seeing, of hearing, of tasting, of smelling, and of feeling,) is that by which we make acquaintance with external objects, and have knowledge of them by means of images and apparitions, or (which is a better expression, as being more general and comprehensive,) by *sentiments* excited in the external organs, through impressions made upon them from objects. (Burthogge 1694, 9–10)

*Reason* or understanding, is a faculty by which we know *external objects*, as well as our own acts, without framing images of them; only by *ideas* or notions. (Burthogge 1694, 10)

*Imagination* is internal sense, or an (after) representation of the images or sentiments (that have been) excited before in the sense. (Burthogge 1694, 10)

Since sensation and imagination are essentially one and the same cognitive power of conceiving by means of “images, or sensible representations” (Burthogge 1694, 10), all human cognitive powers “may be reduced to *two*, to *sense* and *reason*” (Burthogge 1694, 10), the two terms denoting no more than two different forms of conceptualising – in an imaginative or notional way, that is by forming images or by forming notions. Thus the difference between sense and reason comes down, in Burthogge’s view, to the difference between two types of conceiving.

Notions and images produced in the process of conceptualisation (i.e. “conceptions” in the nominal sense of the term) are considered to be the “immediate objects” of conception taken as an act (see above): “conception properly speaking, is of the image, or idea” (Burthogge 1694, 4). However, it must not be forgotten that all the aspects of cognition discussed above, that is its apprehensiveness, cogitativity and conceptivity, form in fact one inseparable whole, so that

conception and cogitation, really are but one act, and consequently, all conceptive are cogitative powers, and cogitative powers conceptive. (Burthogge 1694, 4)

In other words, all three cognitive faculties do “*agree* and concur in this, that they are *conceptive* and *cogitative* [...] powers” (Burthogge 1694, 3). Therefore, not only conceiving, but also cognition as a whole refers to the conceptions (i.e. to notions or imagines) as to its “immediate objects”. On the other hand, as a result of the object-direction given to the knowledge by apprehension, images and notions get a reference also to the ultimate objects of cognition, thereby becoming two conceptually different ways of knowing the external reality (see, however, footnote 58). Apprehension, cogitation, conception – all these operations are inextricably interwoven in the cognitive act, as clearly evidenced by Burthogge’s very first definition of knowledge:

Cogitation is conscious affection; *Conscious affection*, is affection with consciousness of that affection; and by another name is called *knowledge*. Knowledge, as it has a double relation, so it may be considered two ways, to wit, either in reference to the *object*, which is known, and so, properly, it is *apprehension* or conscious perception; or, as it respects the *image* and *idea*, by means of which we do perceive or know that object, and so it may be called *conception*. (Burthogge 1694, 4)

## Modus concipiendi

Since conception is a creation of the mind it must be subject to the conditions imposed on it by the structural-functional properties of the mind. In other words, its mental origin must have some effect, probably profound, on it. Cognition is, therefore, extensively code-termined by the mind. That is Burthogge's most fundamental principle, the implications of which are systematically drawn and discussed throughout his whole epistemology:

Every cogitative faculty, though it is not the sole cause of its own immediate (apparent) object, yet has a share in making it: thus the eye or visive faculty hath a share in making the colours which it is said to see; the ear or auditive power, a share in producing sounds, which yet it is said to hear (..) and there is the same reason for the understanding, that it should have a like share in framing the *primitive notions* under which it takes in and receives objects. (Burthogge 1694, 59)

Thus we have reached a point, when the term "*modus concipiendi*" can appear on stage, for in the technical language of Burthogge's philosophy it refers precisely to that particular aspect of each conception which is determined by the structural and functional properties of the human mind.<sup>7</sup> Since every act of cognition is essentially an act of conceiving or conceptualising, and that conceptualization is always performed in the manner and with the means which are determined by the internal structure of human cognitive faculties, no knowledge is simply a pure grasp of the given data. On the contrary, the mind conceives affection (and through it the external thing) on its own ontological terms. In short, no knowledge is simply given, but is always constructed. Consequently, external reality is never presented to the mind as it is in itself, but under the subjective mode of human conceiving, namely:

---

7 The term as such was not Burthogge's invention. It had been used much earlier, already in medieval philosophy. Furthermore, some closely related terms played an important role in the Aristotelianism of Giacomo Zabarella and the Paduan school (*modus considerandi*, Zabarella 1578) as well as in the anti-Aristotelianism of Arnold Geulincx (*modus cogitandi*, Geulincx 1892a, 1892b). However, within the specific context of Burthogge's epistemology the term gained new meaning and significance.

*modus concipiendi*, a certain particular *manner* of conceiving; a manner of conceiving things that corresponds not to them but only as they are *objects*, not as they are *things*; there being in every conception some thing that is purely *objective*, purely notional; in so much that few, if any, of the ideas which we have of things are properly *pictures*; our conceptions of things no more resembling them in strict propriety, than our words do our conceptions, for which yet they do stand. (Burthogge 1694, 56)

However, this does not mean that the external objects have no role in producing knowledge. In fact, they do, even though it is not easy to assess the exact degree of this participation.<sup>8</sup> Burthogge's complete view on knowledge is, therefore, that it is a result of interaction between the mind and the external thing, in which both of them play their part:

[the external things] by the various impressions that they make upon us [...] do either occasion only, or cause, or (which is most probable) concur unto in causing with our faculties. (Burthogge 1694, 59)

Nevertheless, the creative impact that the mind makes on knowledge is enormous. In fact, the scale of the mind's participation in the creation of knowledge is one of the most original features of Burthogge's approach. Admittedly, also the other philosophers of the time and earlier took some account of the mind's influence on cognition, but only with respect to the formal aspects of cognition. A paradigmatic example in early modern philosophy is the Cartesian theory of intellectual knowledge, in which the mental component of cognition, that is its ideational form of being, is considered to be only a way of internalisation of the object into the mind, thereby not affecting the content of the knowledge being acquired (Descartes 1641, 1644). In contrast to this view, in Burthogge's opinion, an impact made by the mind's activity is universal, affecting the content of knowledge just as intensely as its form, and to such an extent that the external reality becomes unknowable in itself (see next section).

*Modi concipiendi*, much like Kant's pure concepts (Kant 1996, 121–204), have two aspects: one nominal and the other functional. They are conceptions, but unlike other conceptions they are purely mental creations – they are not formed by the standard process of conceptualising affection, that is, with the participation of external things. Nevertheless, they are inherently involved in this process, serving as a primary and inescapable instrument for any conceptualisation (which is the principal reason for describing them as “manners” or *modi* (Burthogge 1694, 56) and consequently providing each standard conception with some “purely notional” content (Burthogge 1694, 56). At the same time, though

---

8 Therefore, Burthogge can say that “the immediate objects of cogitation are external in their *grounds*” (Burthogge 1694, 73, quoted more extensively below).

purely mental in nature, they are not ready-made, innate concepts, but rather (and again in remarkable agreement with the later Kant's view) virtual ones that exist only through and within the particular acts of conceiving.

Furthermore, there are two different types of *modi concipiendi* corresponding to the general division between two basic kinds of conceiving – a sensuous one and an intellectual one. Thus “light and colour” and “sound” serve as model examples for a sensuous manner of conceiving (Burthogge 1694, 57). Burthogge's crucial idea here is to explain the inadequacy of some sensible qualities (as described by Galileo, Descartes, and most famously by Locke) in terms of their mind-dependency – a view that can be seen as one of the key novelties of his epistemology and as a turning point in the whole early modern process of gradual subjectivisation of sense perception. A detailed discussion of this issue exceeds, however, the scope of this article and will be presented in another paper.

As required by the unitary theory of cognitive faculties, an account and explanation of perceptual relativity discussed above is then adequately applied to intellectual cognition, thereby extending the subjectivity of knowledge, hitherto considered specific to sense perception, to cover all kinds of cognition. As a result of this step, Burthogge strikingly anticipates the Kantian theory of categories by postulating the existence of certain mind-dependent universals underlying each intellectual act:

So the *understanding* apprehends not things, or any habitudes or aspects of them, but under *certain notions* that neither have that being in objects, or that being of objects, that they seem to have; but are, in all respects, the very same to the mind or understanding, that colours are to the eye, and sound to the ear. To be more particular, the understanding conceives not any thing but under the notion of an *entity*, and this either a *substance* or an *accident*; under that of a *whole*, or of a *part*; or of a *cause*, or of an *effect*, or the like; and yet all these and the like, are only *entities of reason* conceived within the mind, that have no more of any real true existence without it, than colours have without the eye, or sounds without the ear. (Burthogge 1694, 57–58)

## Burthogge's idealism

Since not only the form, but also the content of all knowledge is partly determined by the cognitive faculties, all what is known, all the cognitive contents and entities, are to be identified as mental products – “phenomena”, “appearances” or “*entia cogitationis*” (Burthogge, 1694, 59–60) – which, as such, cannot be properties or even resemblances of the external things (even though they are partially caused or occasioned by them). Consequently, Burthogge clearly anticipates Kant in claiming that the external world is unknowable in itself, being accessible to the human mind only through the phenomena that the mind itself co-produces – “under the masquerade of *sentiments*”, as Burthogge puts

it (Burthogge 1694, 114).<sup>9</sup> At the same time, there is a fundamental difference between his version of idealism and that advocated, for example, by Berkeley or Arthur Collier. It consists in the explicit claim that there actually is some external. Once again like Kant, Burthogge does not even consider any other possibility:

the immediate objects of cogitation are external in their *grounds*, as well as in appearance, and in truth, are therefore external in appearance, because they are so *really* in their grounds [...]. For *things without us*, are the causes that do excite such images and notions *in us*: in the order of nature, we do *see* a thing so long *as*, and no longer *than*, we keep our Eye upon it; and therefore that we do see it, must come from some *impression* from the thing. (Burthogge 1694, 73)

Thus, ultimately, Burthogge's epistemology turns out to be the idealistic constructivism of a kind that is as Kantian in spirit as it was possible for pre-Kantian philosophy:

It is certain that things to us men are nothing but as they do stand in our *analogy* that is, in plain terms, they are nothing to us but as they are known by us; and as certain, that they stand not in our analogy, nor are known by us, but as they are in our faculties, in our senses, imagination, or mind; and they are not in our faculties, either in their own *realities* [as they be without them], or by way of a true *resemblance* and *representation* [by *picture* and proper representation], but only in respect of certain *appearances* or sentiments [*appearances* and *phaenomena*], which, by the various impressions that they make upon us, they do either occasion only, or cause, or (which is most probable) concur unto in causing with our faculties [...]. In sum, the *immediate objects* of cogitation, as it is exercised by men, are *entia cogitationis*, all phaenomena; appearances that do no more exist without our faculties in the things themselves, than the images that are seen in water, or behind a glass, do really exist in those places where they seem to be. (Burthogge 1694, 59-60; the phrases in brackets come from Burthogge 1678, sect. 9)

It is easy to recognize the crucial steps of Burthogge's argument. First, contrary to most British philosophers of his time, who believed senses to be the supreme means of cognition, Burthogge reintroduced the intellect as an independent source of knowledge.<sup>10</sup> At the same time, by defining sense and reason as apprehensive, cogitative and conceptive powers, he achieved structural and functional unification of the cognitive faculties. Finally, he exploited fully the potential of the early modern discovery of the relativity of sense perception by interpreting it in terms of the mind's activity and applying it effectively to intellectual

---

9 "Sentiment" is frequently used by Burthogge as a synonym for "conception" (see, for example, Burthogge 1694, 24).

10 He was probably influenced in this by Geulincx and other logicians from Leiden (Nuchelmans 1983, 117-20), where he studied and received his medicine doctorate in 1662 (Landes 1920, 254).



cognition. The recipe and ingredients were simple, but the dish was surprisingly innovative and original, even though his contemporaries and the cook himself (as we will see in the next section) hardly noticed it.

## Part Two: Self-Knowledge and Personal Identity

From all that has been said thus far about Burthogge's epistemology, several interesting consequences can be derived for the problems of self-interpretation and self-identity. One general implication of his theory, mentioned above, is that due to the attribution of intentional structure to each cognitive act, neither knowing mind nor object known can be reduced to a pure stream of sense data. On the contrary, they are to be considered ontologically independent from the content of cognition. However, while the distinction presupposed by the very concept of intentional apprehension provides a sufficient basis for differentiation between the general notions of "mind" and "object" (seen as polar opposites of one cognitive act), serious difficulties arise when we consider the identity of any particular subject performing a cognitive act. Basically, three principal strategies to address this issue are available in the post-Cartesian framework of seeing the mind as an ontologically independent entity, and all of them appear to be more or less ineffective when employed in Burthogge's epistemology. The reason for this lies in the limitations of self-knowledge resulting from the fundamental restrictions imposed on cognition by Burthogge's theory. Since these constraints are not placed exclusively on externally directed operations, but apply also to internal cognition, that is to self-knowledge, the conclusion about unknowability of the thing-in-itself, discussed above, refers to the mind just as much as to external objects. Let us examine briefly the implications of this fact.

The first of the above-mentioned strategies for dealing with the problem of personal identity tries to ground it in some extra-mental properties of the subject, preferably material ones. Thus the subject's identity is defined in terms of the relationship of the mind with the outer, material world, and specifically in terms of the observed attachment of the mind to one particular body (referred to as "its own" one) – a concept that is essentially yet another variant of the old Aristotelian view of matter as a principle of individuation. However, as suggested above, Burthogge's epistemology presents insuperable obstacles to the implementation of such a solution. As, according to his view, the mind has no access to the external world in itself, and consequently human knowledge is confined within the realm of mental phenomena (which the mind itself co-produces), the subject's identity cannot be defined in terms of the relationship to the reality about which it knows nothing except that it exists. Matter simply cannot be considered the principle of individuation when there is no certainty that there is any such thing as matter at all. Admittedly, the concept of matter, being (as any other concept) partially caused or occasioned by external things, is



somehow grounded in the external world. Nevertheless, it presents the outer world as it is seen through the subjective “glasses” of human conceiving. In other words, what the mind experiences as matter is, on Burthogge’s principles, merely a mental construct, formed by combining the intellectual conceptions of substance and accidents with no less conceptual sensible properties of extension, motion, etc. Therefore, to explain the subject’s identity in material terms would be essentially the same as to connect it with a certain selected set of mental objects – which, in fact, would mean changing the whole strategy for dealing with the problem in question.

The serious problem in interpreting Burthogge’s philosophy is, however, that most often he does explain personal identity in terms of individuation by matter:

were spirits absolutely pure and simple, without any concretion of matter, there could be no distinction among them as to *individuals*, as well as none in relation to kinds. For since all individuation (except only that of the central pure mind) is *numerication*, and all numerication arises from division, and division has no place but in Matter, or in things by means of matter. It is evident that there can be no distinction of spirits as to individuation, if there be no ingredience of matter in their making [...] in short, we may observe *in our selves*, (that mind as I have noted before) is *individuated by matter*. (Burthogge 1694, 167–168)

At the same time, another view – one probably more consistent with his general nominalist standpoint (Burthogge 1694, 60–61) – can also be found in his writings, namely that spirits need not to be individuated by any special principle, but are individual by their nature (Thiel 2011, 73):

they seem to me to come nearest to the truth, who do affirm, that a singular or individual becomes so, not by any distinct principle of individuation, but immediately *per se*, and in that, that it is in being. (Burthogge 1694, 270)

While inconsistency between the above two statements is only apparent, as has been shown by Thiel (Thiel 2011, 73), there is another, much deeper, inconsistency between his free considerations about reality in itself and his epistemological principles, as they had been laid down in *Essay*. This inconsistency reflects a more general duality running throughout all his work. While the epistemological theory presented in the first chapters of his essay anticipates, as we have seen, Kantian idealism, the remaining part of his deliberations represents almost entirely a dogmatic point of view – to put it in Kantian terms (Kant 1996, 34). Although sometimes Burthogge seems to recognise this inconsistency by warning that his discussion is made only in the “notional way” (Burthogge 1694, 96, 106), through the whole *Essay*, as if by a gestalt switch, his standpoint changes from Kantian-type idealism to Cartesian-type dualism. This might explain why his revolution has remained unnoticed – it was because it had not been noticed by Burthogge himself.

The second of the above-mentioned strategies for solving the problem of personal identity is to define it in terms of intrinsic properties of the mind. At first glance, this method seems more promising than the first one. Considered from the purely theoretical point of view, Burthogge's epistemology, at least in its initial form presented in the first half of *Essay*, belongs more to the modern, post-Kantian, philosophy of consciousness, than to the tradition of early modern dualism. This affiliation is confirmed by three original features of his approach.

First, as we have seen, every human cognition presupposes some elementary form of consciousness or awareness. Consequently, each cognitive act of the mind can accurately be considered a conscious or, as Burthogge's puts it, a "cogitative" act. Accordingly, "cogitation" is to be treated as the first and the most general synonym for knowledge. Secondly, and more importantly, Burthogge makes every effort to achieve what can be called the extensive mentalisation of the cognitive powers and acts. Specifically, he firmly and deliberately rejects any attempts to explain sense perception in purely materialist or mechanistic terms:

sense and imagination, as well as the understanding and reason, are mental and spiritual, not merely mechanick and material powers. (Burthogge 1694, 8–9)

The sense, sensuous acts and sensible qualities are, therefore, mental phenomena, and they are so, as we have seen, for ontological, structural reasons. Finally, a remarkable idea of structural and functional uniformity among all human cognitive faculties and acts enables Burthogge to provide a unified theory of consciousness, which is similar, in its crucial points, to modern phenomenology rather than to the empiricist or rationalist philosophy of his time.

Nonetheless, the idea of explaining personal identity in terms of the structural properties of the mind also faces serious obstacles, mostly for the same reasons as the first strategy did. Since self-knowledge is no exception to the structural-functional principles imposed on cognition, the mind can also be known only under the subjective conditions of human conceiving. It is, therefore, just as unknowable in itself as the external things are. Accordingly, even such a general definition of the mind as "cogitative, thinking, or perceiving substance" (Burthogge 1694, 106) must be tempered by the reservation that

we have not any real immediate conception of [substance], but only a notional. Or (to speak more plainly, according to the principles laid before) substance as such, is not a thing conceived just as it is in its own reality, but a thing conceived under a certain notion. (Burthogge 1694, 96–97)

All in all, the mind can be self-interpreted at most indirectly – as it manifests itself by its faculties, powers or acts. In other words, the situation of Burthogge's subject is very similar to that of Kant – in between the two noumenal realities, one external and the other internal.

The last and apparently the most fruitful strategy for dealing with the question of personal identity in accordance with the basic principles of Burthogge's philosophy is the

one suggested by Locke in his own discussion of the problem. On this approach, suitably adapted for the purposes of Burthogge's theory, the subject's identity would be determined by the specific set of conceptions (immediate objects of cognition) appearing in one's mind. A great advantage this solution has over the other two methods is that in principal it does not require any cognitive access to the external reality. When applied to Burthogge's epistemology it has one major disadvantage, though (aside from the general defects discussed extensively in the literature). When no knowledge of the subject in itself is available, the identity defined in this way may be completely false. Thus it is a sense of identity rather than identity as such that is examined with this method. In fact, one might just as well argue – as, indeed, it was argued by some of Kant's successors headed by Hegel – that, instead of many internal and external substances, there is, in reality, only one substantial mind, thereby opening the door to the monistic idea of absolute consciousness; a theoretical shift that fortunately (or unfortunately, depending on one's perspective) did not take place in British philosophy, because of the minimal impact Burthogge's ideas had on it.

From what has been said so far, it is clear that Burthogge's epistemology suffers from substantially the same disadvantages as those of Kant. This similarity certainly deserves a close examination, which obviously cannot be conducted here. However, one general historical conclusion can be drawn already at this point. It concerns Kant's idealism. Rather than viewing it as a unique and isolated theory, we should consider it the most significant representative of the whole class of constructivist-idealistic doctrines that emerged in response to the problems created by Cartesianism. The main purpose of this paper was to show how the writings of the lesser-known English philosopher can shed some light on the way to achieve this goal.

## Acknowledgment

This work was supported by the National Science Centre, Poland under grant number 2015/19/D/HS1/01013.

## Bibliography

### Primary Sources

- Arnault, Antoine and Nicole, Pierre. 1662. *La Logique ou L'Art de Penser*. Paris: Charles Savreux.
- Arnault, Antoine, and Nicole, Pierre. 1674. *Logica, sive Ars Cogitandi*. London: Typis Andr. Clark, Impensis Joh. Martyn sub Signo Campanae in D. Pauli Coemeterio; Et Ed. Story Bibliopolae Cantabrigiensis.
- Arnault, Antoine, and Nicole, Pierre. 1685. *Logic; or, the Art of Thinking*. London: T. B. for H. Sawbridge, at the Bible on Ludgate-hill.

- Burthogge, Richard. 1678. *Organum Vetus & Novum. Or a Discourse of Reason and Truth. Wherein The Natural Logick common to Mankind is briefly and plainly described.* London: Printed for Sam. Crouch, at the Princes Arms a Corner-shop of Popeshead ally in Cornhil.
- Burthogge, Richard. 1694. *An Essay upon Reason, and the Nature of Spirits.* London: Printed for John Dunton at the Raven in the Poultry.
- Descartes, René. 1641. *Meditationes de prima philosophia.* Paris: Michael Soly.
- Descartes, René. 1644. *Principia philosophiae.* Amsterdam: Lodewijk Elsevier.
- Geulincx, Arnold. 1892a. “Metaphysica ad mentem Peripateticam.” In *Arnoldi Geulincx Antverpiensis Opera Philosophica*, vol. 2, edited by J. P. N. Land, 199–265. The Hague: Martin Nijhoff.
- Geulincx, Arnold. 1892b. “Metaphysica vera.” In *Arnoldi Geulincx Antverpiensis Opera Philosophica*, vol. 2, edited by J. P. N. Land, 139–198. The Hague: Martin Nijhoff.
- Kant, Immanuel. 1996. *Critique of Pure Reason.* Translated by Werner S. Pluhar. Indianapolis/Cambridge: Hackett.
- Zabarella, Giacomo. 1578. *Opera logica.* Venice: Apud Paulum Meietum Bibliopolam Patavinum.

## Secondary Sources

- Ayers, Michael Richard. 2005. “Richard Burthogge and the Origins of Modern Conceptualism.” In *Analytic Philosophy and History of Philosophy*, edited by Tom Sorell and G. A. J. Rogers, 179–200. Oxford: Clarendon Press.
- Cassirer, Ernst. 1922. *Das Erkenntnisproblem in der Philosophie und Wissenschaft der neueren Zeit. Erster Band.* Berlin: Verlag Bruno Cassirer.
- Grünbaum, Jakob. 1939. *Die Philosophie Richard Burthogges (1637–1698).* Bern: Buchdruckerei J. Kleiner.
- Landes, Margaret W. 1920. “Richard Burthogge, His Life and His Place in the History of Philosophy.” *The Monist* 30(2): 253–66.
- Landes, Margaret W. 1921. “Introduction.” In *The Philosophical Writings of Richard Burthogge*, edited by Margaret W. Landes, xi–xxiv. Chicago–London: The Open Court Publishing Company.
- Lyon, Georges. 1888. *L’Idéalisme en Angleterre au XVIII<sup>e</sup> siècle.* Paris: Ancienne Librairie Germer Baillière et Cie Félix Alcan, éditeur.
- Nuchelmans, Gabriel. 1983. *Judgment and Proposition. From Descartes to Kant.* Amsterdam – Oxford – New York: North-Holland Publishing Company.
- Rutherford, Donald. 2006. “Introduction.” In *The Cambridge Companion to Early Modern Philosophy*, edited by Donald Rutherford, vols 1–9. Cambridge – New York: Cambridge University Press.
- Sgarbi, Marco. 2012. “L’epistemologia di Richard Burthogge.” *Rivista di storia della filosofia* 3: 493–521. <https://doi.org/10.3280/SF2012-003003>
- Šoć, Andrija. 2012. “Locke’s Anticipation of Idealism.” *Theoria* 55(3): 99–112. <https://doi.org/10.2298/THEO1203099S>
- Thiel, Udo. 2011. *The Early Modern Subject.* Oxford – New York: Oxford University Press.

# Knowing Me, Knowing You: Berkeley on Self-Knowledge and Other Spirits

It is well-established (and explicit in the work itself) that one of Berkeley's chief aims in *A Treatise Concerning the Principles of Human Knowledge* is to collapse the distinction between what we perceive and what exists (and, more broadly, between perception and existence).<sup>1</sup> This distinction, Berkeley believes, is a product of irresponsible and pernicious abstract thought. Abstract thought, Berkeley maintains, strays too far from what is experienced, what is conceivable and, as he seeks to demonstrate throughout the *Principles* and the *Three Dialogues*, what is possible.<sup>2</sup> As he puts it in PHK §5, there is no better example of abstract thought than, “to distinguish the existence of sensible objects from their being perceived, so as to conceive them existing unperceived”. As Berkeley sees it, the notion of the existence of matter, material substance or material objects requires this distinction be upheld, and thus he argues that “materialism” (the broad view on which material substance exists, either exclusively or as well as spiritual substance) ought to be rejected. Berkeley's deflationary approach to our perception of the external world leads him to the famous claim that when it comes to the sensible things that make up the world around us – “Their *esse* is *percipi*” (PHK §3). For Berkeley, there is no distinction to be drawn between the existence of a thing and its being perceived. To enquire separately into (i) the existence or nature of things, and (ii) the manner in which we perceive them (as materialists do) is to follow an erroneous line of enquiry thrown up only by abstract thought.

My aim in this paper will be to demonstrate that matters of ontology, perception and immediate knowledge, for Berkeley, ought not to be enquired into separately – even in the

---

1 In what follows, references are to the final 1734 edition of the *Principles*.

2 For Berkeley's remarks on the dangers of abstract thought, see PHK §97, §99; AMP 7.15. For an insightful discussion of the issues surrounding the relation between conceivability and possibility in Berkeley's thought, see Holden 2017.

cases of self-knowledge and knowledge of other minds.<sup>3</sup> As I will demonstrate, identifying the nature of either the perceiver or the perceived is, for Berkeley, one and the same with determining what exists. I will focus on the nature of self-knowledge; knowledge of oneself as a spirit. Berkeley's strict commitment to the view that it is impossible to have an idea of a spirit means that spirits cannot be *perceived* (PHK §27). In other words, a spirit (which is active) cannot be a passive object to an act of perception. However, I will demonstrate that Berkeley is nonetheless equipped with an account of self-knowledge – of both the existence and nature of one's own spirit – which ought nonetheless to be understood in terms of perception (more broadly).

My reading of Berkeley's account of self-knowledge is motivated by his claim that it is only by virtue of having immediate knowledge of the self – in other words, one's *own* spirit – that we can gain knowledge of other spirits. Two claims, in particular, make it clear that, for Berkeley, self-knowledge is something we *use* in order to gain knowledge of other spirits. In the *Three Dialogues*, Berkeley explains that;

My own mind and ideas I have an immediate knowledge of; and *by help of these*, do mediately apprehend the possibility of the existence of other spirits and ideas. (DHP 231–2 [my emphasis])

While in the *Principles* he claims that:

we know other spirits *by means of our own soul*, which is in that sense the image or idea of them. (PHK §140 [my emphasis])

The first claim indicates that Berkeley thinks all spirits are alike (somehow or another); he argues that the immediate knowledge I have of myself as spirit is a tool used to gain knowledge of the existence of others. The second implies similarly that my own spirit is key to understanding what others are like. Together, these claims indicate that knowledge of both the existence and nature of other minds depends upon the immediate knowledge we have of our own mind. As such, Berkeley places a significant epistemological burden on his account of *self-knowledge*. In fact, his accounts of self-knowledge and knowledge of other spirits stand or fall together. Unless Berkeley can provide an adequate account of the former, he *de facto* (by his own principles) fails to provide an account of the latter.

In making this claim, Berkeley cannot mean to say that my own spirit is literally an “image or idea” of other spirits – this is ruled out by Berkeley's strict epistemology (PHK §27). One of the central tenets of Berkeley's immaterialist system is that there are two kinds of existent: those which exist by virtue of *being perceived* (ideas) and those which exist by virtue of *perceiving* (spirits). Berkeley maintains that since these two kinds of existent are

---

<sup>3</sup> In what follows, “self-knowledge” refers to the immediate knowledge we have of ourselves *as spirits*. Like Berkeley, I use the terms “spirit” and “mind” interchangeably (see PHK §2).

entirely heterogeneous, a passive idea could never represent an active spirit (PHK §142).<sup>4</sup> Yet, in the opening passage of the *Principles*, Berkeley suggests that we gain knowledge by means of our ideas (PHK §1). In what sense, then, can we gain knowledge of either our own or other spirits? If Berkeley does not have a clear story of how we gain self-knowledge, then (by his own reasoning) he is lacking a foundation on which to base his epistemology of other minds. What makes this concern particularly pressing (bearing in mind Berkeley's theistic concerns)<sup>5</sup> is the fact that, as Berkeley makes clear, this goes for not only other human minds but also the mind of God (PHK §147).<sup>6</sup> My aim will be to show that Berkeley's account is not undermined by such concerns and that he *is* equipped to provide an account of self-knowledge – which ought to be construed in terms of perception.

The structure of my argument is as follows. In the first section I will emphasise the extent to which Berkeley moves away from what I characterise as a Cartesian account of the self and self-knowledge. In the second section I will argue that Berkeley's account of self-knowledge provides us with immediate knowledge (or “immediate data”) that can plausibly be used in order to gain knowledge of other spirits. Finally, I will argue that while Berkeley's account of self-knowledge is not straightforwardly a matter of *what is perceived*, it is nonetheless grounded in perceptual activity.

## Berkeley's Rejection of Cartesian Self-Knowledge

In this section I will emphasise Berkeley's move away from “abstract” or “absolute” conceptions of the self and *inferential* accounts of self-knowledge – particularly the Cartesian account.<sup>7</sup> Above all, in order for Berkeley to assert that self-knowledge is the kind of knowledge that can be employed in gaining knowledge of other minds, he must establish how the self, or one's own spirit, can be an object of immediate knowledge at all.

---

4 This rests on Berkeley's acceptance of what we might call “the resemblance thesis” which dictates that in order for one thing to represent another, the two things must *resemble* one another. See MI §12; PHK §§56–7.

5 Both PHK and DHP are anti-sceptical and anti-*atheistic* works. The front matter to DHP sets it “In opposition to sceptics and atheists”.

6 I will not discuss Berkeley's account of our knowledge of God. This is a contested issue in its own right. Literature surrounding this issue focuses primarily on Berkeley's account of “divine analogy”. For Berkeley's most thorough discussion of divine analogy, see AMP 4.16–22.

7 Stephen H. Daniel (2018) has recently provided a more straightforwardly metaphysical account of Berkeley's move away from the Cartesian account of spirit (my own emphasis is largely epistemological). Daniel notes that many commentators take it for granted that Berkeley adopts a largely Cartesian view of spiritual substance and selfhood (660, footnote 2). It will become clear that I, like Daniel, reject such an assumption.



Throughout this section, I will contrast Berkeley's account of self-knowledge with an inferential account which I attribute to Descartes. In doing so, it might be argued, I have failed to acknowledge Descartes' claim in *Replies to the Second Set of Objections* that the "cogito" is not an argument or inference but rather a "primary notion" (CSM 2: 100). Here Descartes writes:

When someone says "I am thinking, therefore I am, or 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.

This, as Descartes points out, is necessary if he is to avoid positing there is a general principle at work – "Everything which thinks is, or exists" – from which the particular claim "I exist" is derived. Were such a general principle at work, Descartes could not consistently maintain that the "cogito" is a first principle of knowledge (i.e. the "firm and immovable point" that he wants it to be (CSM 2:16). In this way, Descartes establishes the certainty with which we know the existence of the self. What's more, Descartes distances our knowledge of the *existence* of the self from our knowledge of *what that self is like*; once the "cogito" is established he *then* goes on to examine what it is that "I" refers to (CSM 2:19). As we will see, Berkeley does not acknowledge any such distinction. It will become clear that this is one of the key differences between a Cartesian and a Berkeleian account of the self.

Regardless of whether Descartes can plausibly argue that "I am thinking, therefore I am" is prior to a more general principle, it is clear that some kind of inference is at work in his account of self-knowledge – of how we gain knowledge of ourselves as thinking things. The Cartesian account of self-knowledge is an inferential one – and Descartes' claims (outlined above) should not deter us from accepting that. The way that Descartes *phrases* his self-knowledge claims indicates that there is a step from (i) knowledge of oneself as thinking, to (ii) knowledge of oneself *as oneself* (i.e. *a thinking thing*). Even if, for Descartes, these two kinds of knowledge are inseparable in the sense that once I know (i) I cannot help but know (ii), it becomes clear that knowledge of oneself as thinking is necessary *before* one can have knowledge of oneself as a thinking thing (we might say that the latter "bootstraps" on the former). If no such inference were required, then Descartes' use of the term "therefore" would be redundant (CSM 1:194–95, 1:127, 2:100). Thus, even if no argument or syllogism is required in order to reach the conclusion that "I exist", I will take it that gaining knowledge of oneself does, for Descartes, involve a move from the existence of thought (or modes of thought) to the existence of thinker. This is at least consistent with Descartes' claim that the mind comes to realise, "that it is impossible that [it] should think without existing" (CSM 2:100).<sup>8</sup>

---

8 See also Descartes' replies to Hobbes's second and third objections where he makes a number of such claims including: "I do not deny that I, who am thinking, am distinct from my thought, in the way in which a thing is distinct from a mode [...]. And when I add, 'Which of them can be said to be separate from myself?', I simply mean that all these modes of thinking inhere in me." CSM 2:123–4. See also CSM 1:196, 1:210, 1:215.



As far as Berkeley is concerned, what is straightforwardly clear is that self-knowledge is not the same kind of immediate knowledge we have of our ideas; because spirit is neither the kind of thing that can be represented by ideas nor the kind of thing that can be perceived (Bettcher 2007, 41–42).<sup>9</sup> In fact, as Berkeley makes clear in PHK §142, spirits and ideas are so unlike one another that when we use phrases like “they exist” or “they are known”, the terms involved are used in totally different ways. As Kenneth Pearce puts it, “there are [for Berkeley] two distinct senses of ‘know’, and in one of these senses it is correct to say ‘spirits are known’ and in the other it is not” (Pearce 2017, 126).<sup>10</sup> I cannot have ideational knowledge of my own spirit, but I can have a different (“notional”) kind of knowledge of it (PHK §27, §142).<sup>11</sup> So what exactly does this other kind of knowledge – immediate knowledge of the self – consist in? What does it mean to have non-ideational but immediate knowledge of oneself as a spirit?

If we grant that Descartes’ account is an inferential one, I take it as uncontroversial to draw a sharp distinction between the Cartesian account of self-knowledge and Berkeley’s own account. Although commentators disagree as to what Berkeley’s view on the nature of spirit and its relationship with the ideas it perceives is, it is clear he does not subscribe to an inferential account of self-knowledge.<sup>12</sup> It will be helpful to emphasise the extent to which Berkeley breaks with the outlined Cartesian approach for the sake of identifying what is distinctive about his own account.

In *Principles of Philosophy*, Descartes makes the claim, early on, that

it is a contradiction to suppose that what thinks does not, at the very same time when it is thinking, exist. Accordingly, this piece of knowledge – *I am thinking, therefore I exist* – is the first and most certain of all to occur to anyone who philosophizes in an orderly way. (CSM 1:195)

In this case, Descartes appears to infer from an implicit principle that “what is thinking must also exist” that something (an “I”) exists in order to explain the fact that such thinking is taking place.<sup>13</sup> Similarly, in the second meditation, we are told that:

---

9 Again, this requires acceptance of the “resemblance thesis”. In §8 of the *Principles*, Berkeley states that “an idea can be like nothing but an idea” (‘the Likeness Principle’). This, in turn, means that an idea cannot resemble – and therefore cannot *represent* – a spirit (since a spirit is not an idea) (PHK §27).

10 See also Winkler 2011, 227–9; Daniel 2018, 660.

11 Berkeley’s use of the term “notion” is a point of contention amongst commentators, but that discussion is beyond the scope of my current concerns. The most developed account of Berkeley’s “doctrine of notions” is Flage 1987.

12 Stephen H. Daniel denies that Berkeley saw spirits as subjects in a perceptual relationship but rather as “the activity or principle whereby objects are identified in relation to one another” (2018, 660; see also 2008, 204–5). On the other hand, Bettcher and Pearce both maintain that Berkeley’s spirits are subjects that perform acts of perception (Bettcher 2007, 42–54; Pearce 2017, 132).

13 As such, Descartes appears to undermine his claim in *Replies to the Second Set of Objections* that there is no general principle – “what is thinking must also exist” – at work when I come to know the existence of my mind.

after considering everything very thoroughly, I must finally conclude that this proposition, *I am, I exist*, is necessarily true whenever it is put forward by me or conceived in my mind. (CSM 25:17)

Here, Descartes' claim is that the very act of thinking provides demonstrative evidence of the existence of a thinking thing; from the fact that I am thinking I know, with certainty, that a thinker exists. While Descartes is explicit in claiming that I am *inseparable* from my thoughts and suggests I exist "as long as I am thinking" (CSM 2:18), the distinction between agent and action is nonetheless clear. In this way, Descartes posits the existence of both (i) thought and (ii) thinker; different modes of thought and the agent or subject who performs them – and distinguishes between the two. Consider, for example, the following claim.

The fact that it is I who am doubting and understanding and willing is so evident that I see no way of making it any clearer. (CSM 2:19)

There is obviously an implicit distinction between the subject and its actions at work, detectable from the fact that "I" is distinguishable from doubting, understanding and willing. This is further corroborated by remarks in *Principles of Philosophy* where, for example, Descartes writes: "thought alone belongs to *it* [my mind]" (CSM 1:195). Thought cannot, strictly, belong to itself, so there must be some subject (my mind) to which it belongs. In metaphysical terms, this assumption is grounded in Descartes' adoption of a substance-mode ontology; by identifying different modes of thought (ways of thinking) it is possible to infer the existence of that substance in which they inhere (my mind).<sup>14</sup>

In fact, that Descartes is even able to ask the question "What then am I? [...] What is that?" (CSM 2:19) demonstrates that the existence of the self can be known separately from its actions. Indeed, even though the existence of thought is key to inferring the existence of the mind (knowing that "*cogito ergo sum*"), Descartes' view is that it is possible to gain knowledge of the self *before* becoming clear regarding what type of thing it is (a thinking thing) or the types of actions it performs (such as doubting, understanding, willing).<sup>15</sup>

Berkeley, on the other hand, insists that such an inference does not – and indeed could not – provide knowledge of anything beyond those mental operations (in Cartesian terms, those modes of thought). Berkeley rejects the claim that knowledge of the existence of the self is prior to knowledge of the *nature* of the self, just as he rejects that knowledge of the existence of sensible objects is distinct from knowledge of their nature (PHK §5).

---

14 Daniel argues that Berkeley rejects this broadly Aristotelian subject-mode ontology of mental substance entirely. See 2008, 674, 679–80.

15 See Thiel 2011, 38.

Like Descartes, he maintains that “the soul always thinks” (PHK §98).<sup>16</sup> However, Berkeley goes further, explaining:

[Therefore] whoever shall go about to divide in his thoughts, or abstract the existence of a spirit from its cogitation, will, I believe, find it no easy task.

For Berkeley, there is nothing to be gained by attempting to understand the self as abstracted from those actions by which we come to know its existence. If the soul always thinks, then what knowledge can we possibly hope to have of it aside from its cogitation? (NB 652) He compares the situation to attempts to conceive of “extension” or “motion” in abstract. In both such cases, he argues, once we abstract away the qualities by which we come to know “extension” and “motion”, there is nothing left for us to conceive of – and thus no further knowledge to be gained (PHK §99). To enquire into such things or to claim that our knowledge of their existence is distinct from our knowledge of their particular qualities, he insists, is merely “barren speculation” (my emphasis) – to look for knowledge where there is none available (PHK §156; AMP 7.15).

The parallels between self-knowledge (knowledge of one’s own spirit) and attempts to gain knowledge of sensible things (knowledge of one’s ideas) in abstract are important. At the beginning of the *Principles*, where Berkeley is reacting to the “strangely prevailing [i.e. widespread] opinion” that sensible objects exist independently of the mind, he asks whether there can be

a nicer strain of abstraction than to distinguish the existence of sensible objects from their being perceived, so as to conceive them existing unperceived? [...] [I]s it possible to separate, even in thought, any of these from perception? For my part I might as easily divide a thing from itself. (PHK §§4–5)

The approach is clearly rhetorical and Berkeley’s point is that when it comes to sensible objects *to exist is one and the same as to be perceived*; the one cannot be divided from the other. After all, a sensible thing’s *esse* is *percipi* (PHK §3). This is confirmed in PHK §17 where Berkeley provides us with a warning against attempts to conceive of existence in abstract: “The general idea of being appears to me the most abstract and incomprehensible of all other.” As we saw previously, spirits and ideas are neither known nor exist in the same way; they are as categorically different as sounds and colours are to the senses (PHK §142). We could no more come across a sensible spirit than we could a square circle. Nonetheless, Berkeley’s remarks about what the existence of a spirit consists in make two things clear: (i) that existence (in a wider sense than just sensible existence) is always tied to perception, and (ii) that Berkeleian self-knowledge is not inferential.

This is supported by various remarks from entries in Berkeley’s *Notebooks*, where he is quite explicit about the relationship between existence and perception. In entry 646, he

---

<sup>16</sup> See CSM 1:127.

writes: “Existence not perceivable without perception or volition not distinguish’d therefrom.” As before, the claim is that once we abstract the existence of things from the acts of perceiving there is nothing left for us to perceive.<sup>17</sup> That is, while ideas exist by virtue of *being perceived*, spirits exist by virtue of *perceiving*. This is likewise suggested by a further entry where Berkeley maintains that “Existence is percipi or percipere [...] or velle i.e. agere”; to be is to be perceived, to perceive, or to will (or act) (NB 429).<sup>18</sup> To exist, therefore, is either to be passively perceived (to be an idea) or to actively perceive in one way or another (to be a spirit). This is confirmed in PHK §89, where Berkeley writes:

*Thing or being* is the most general name of all, it comprehends under it two kinds entirely distinct and heterogeneous, and which have nothing common but the name, to wit, *spirits* and *ideas*.

That is, to be a “thing” or a “being” – in other words, to exist – is to be either a spirit or an idea. These are the two kinds of “thing” or “being” in existence, and the terms ought not to be applied to anything above and beyond them.

Although “exist” means different things when applied to spirits and ideas, it can nonetheless, in both cases, be ultimately cashed out in terms of *perception*. In the *Principles*, Berkeley explains that “spirit is one simple, undivided, active being” and later describes it as “that which acts [...] that active principle of motion and change of ideas” (PHK §27). He explains that “understanding” is what we call the spirit when it perceives, and that “will” is what we call it when it either produces ideas or “operates about them.” Furthermore, he argues that it is not possible, even conceptually, to distinguish these “ways of perceiving” from the perceiver (that is, an individual spirit) itself. In doing so he denies that, aside from the awareness we have of “willing” and “understanding”, there is

a third idea of substance or being in general, with a relative notion of its supporting or being *the subject* of the aforesaid powers, which is signified by the name “soul” or “spirit”. (PHK §27 [my emphasis])<sup>19</sup>

This is significant for my current purposes, since it indicates that to enquire into what terms like “soul” or “spirit” signify, aside from their operations, is, again, “barren speculation”. The view that we can gain knowledge of the “subject”, “soul” or “spirit” abstracted from its actions is “what some hold” but is refutable on the basis of both experience and the limits of our conceptual abilities,<sup>20</sup> and thus ought to be rejected. Although this argument need not be read straightforwardly as a rejection of the Cartesian view, it is nonetheless

---

17 Berkeley’s remarks in PHK §142 demonstrate that volition or “willing” is an act of perception.

18 See also NB 674.

19 See also NB 580–1.

20 For Berkeley’s account of what we can and cannot conceive, see the *Principles* Introduction, especially §10 and §22.

sufficient to differentiate Berkeley's account of self-knowledge from that position.<sup>21</sup> Unlike Descartes, Berkeley quite clearly denies that we could *first* be certain that "I" exist before *then* asking what "I am" (CSM 2:19).

These metaphysical claims make it clear that the being, or existence, of a spirit consists in its perceptual activities; its *esse* is *percipere*. The epistemological ramifications of this view are that in order to gain knowledge of spirit, it is sufficient that we gain knowledge of those activities. Berkeley reiterates this view in the *Principles* when he attacks attempts to "abstract the existence of a spirit from its cogitation" (PHK §98). As such, so long as a self-knowledge is possible, it must consist in immediate knowledge of our perceptual *activities*.

## Self-Knowledge as Immediate Data

I previously stated that in order for self-knowledge to be used as a means for gaining knowledge of other spirits we must become clear about what it means to have immediate knowledge of "my own spirit". Berkeley clearly states that "we know other spirits by means of our own soul" – so he *must* have an account of self-knowledge in order to form the basis of his wider epistemology of mind.<sup>22</sup>

We have established that Berkeley does not see self-knowledge as something that can be abstracted from knowledge of one's own perceptual activity. To know the self is to know the distinct features of the mind or the different ways in which it perceives (and since for a mind to exist just is for it to perceive: the different ways in which it exists). It is not possible to conceive of the self *in absolute*; that is, as separate from the *means by which we come to know it* (as the Cartesian would have us do). For Berkeley, we can no more come to know our own mind beyond our perceptual activities, than we can come to know a sensible object abstracted from its sensible qualities. In order to understand this view, it is important to bear in mind that Berkeley does *not* see ideas as operations of the mind; they are distinct and passive entities that are *acted upon* (i.e. perceived) by the spirit as it is called "will" and "understanding" (PHK §27). If self-knowledge is reducible to knowledge of our perceptual acts, then it is reducible to knowledge of *what it is like* to perceive certain ideas (and not reducible to the ideas themselves, for the reasons outlined above).

---

21 For an explanation of why Berkeley's claims undermine both Cartesian and Lockean accounts of spiritual substance, see Daniel 2008; 2018.

22 One reviewer pointed out to me that it may well follow from the fact that, for Berkeley, we cannot gain knowledge of the self beyond its (perceptual) activity that Berkeley does not, in fact, have an account of *self-knowledge* at all. That is to say, perhaps there is, for Berkeley, a self beyond its perceptual activities, but we just cannot gain knowledge of it. However, Berkeley's claim that we know other spirits by virtue of the knowledge we have of our own seems to preclude this possibility (see PHK §140; DHP 231). The claim in DHP 231 that "*My own mind* and ideas I have an immediate knowledge of" is especially indicative of this.

In order to *use* the knowledge we have of our own spirit as a step towards knowledge of other spirits, Berkeley must explain how self-knowledge provides us with what we might call “immediate data”.<sup>23</sup> In other words, it must be the case that, in being aware of one’s self as a spirit, one gains immediate knowledge of the kind that can provide a step towards that which we do not (immediately) know. It must provide us with data we can use as a *sign* – or, as Berkeley puts it, an “image or idea” (DHP 231–2). What needs clarifying is the nature of the “data” we are drawing on when we gain knowledge of other minds.

In the third *Dialogue*, Berkeley lays out the parallels between immediate knowledge of sensible things and immediate knowledge of spirit. Though they exist in categorically different ways, which means our knowledge of them is likewise categorically different, there nonetheless *are* parallels to be drawn. He writes:

I own I have properly no idea, either of God or any other spirit; for these being active, cannot be represented by things perfectly inert, as our ideas are [see PHK §142]. I do nevertheless know that I, who am a spirit or thinking substance, exist as certainly as I know my ideas exist. (DHP 231)

This latter claim is especially important: our self-knowledge is *as certain* as knowledge of our ideas.<sup>24</sup> This is significant in light of Berkeley’s claims about what it means to immediately perceive our ideas. For example, he maintains that “I cannot be deceived in thinking I have an idea which I have not” (PI §22). I also cannot be mistaken as to their nature; as Hylas puts it, “I know them perfectly” (DHP 206).<sup>25</sup> Berkeley also claims (this time via Philonous) that

I know what I mean by the terms “I” and “myself”; and I know this immediately, or intuitively, though I do not perceive it as I perceive a triangle, colour, or a sound. (DHP 232)

Here Berkeley reiterates the claim that although I do not *perceive* whatever it is that is signified by the terms “I” and “myself”, my knowledge thereof is no less immediate than the knowledge I have of my ideas. Again, this is an *un*-Cartesian move to make: no inference is required in order to gain knowledge of ideas and thus, in light of Berkeley’s parity claims, no inference ought to be required in order to gain self-knowledge. So what can self-knowledge, without inference, consist in?

---

23 This terminology is borrowed from Russell 2001, 20. Russell uses the term to refer to Berkeleyian *ideas* specifically, while I am using it to refer to perceptual data more broadly (i.e. I use it to refer to *perceptual acts* as well as *perceptual objects*).

24 Both Adams and Winkler go even further, arguing that self-knowledge is, for Berkeley, infallible. See Adams 1973, 47–69; Winkler 2011, 239–40.

25 In this instance, Hylas (the materialist) reflects Berkeley’s own view, as he is led to this conclusion by Philonous via a series of leading questions.

In the domain of ideas (in cases of both sensation and introspection)<sup>26</sup> immediate knowledge, for Berkeley, means immediate experience of determinate qualities, while mediate knowledge involves using this immediate knowledge – these ideas – as signs in order to make more general or universal claims about the world (PI §12). Ideas constitute “immediate data” insofar as they are determinate in nature and are not known inferentially. The *esse* of ideas is *percipi*, while the *esse* of spirits is *percipere* (or *velle*, i.e. *agere*). Thus, in either case, gaining immediate data ought to involve perception. Perception, for Berkeley, can be construed in two (non-abstract ways): in terms of either what is perceived (ideas) or acts of perceiving (which are attributed to spirits). Perceptual *data* can either take the form of perceived ideas (which provide us with knowledge about themselves) or else knowledge of the acts of perceiving (whatever form that might take).<sup>27</sup> The question, then, is whether Berkeley provides an account of immediate perceptual *data* concerning the self. Of course, this perceptual data won’t take the form of ideas (i.e. things perceived), but perception, for Berkeley, provides us with knowledge of more than just what is perceived.

So what form does immediate perceptual data about the self take for Berkeley? My claim is that it is in outlining the difference between the spirit as it is called “willing” and the spirit as it is called “understanding” – i.e. the different *ways in which we perceive ideas* – that Berkeley gives an account of immediate data, and thus immediate knowledge, of the self. Knowledge of the self as spirit, for Berkeley, does not take the form of ideas but is constituted by the *experience* of perceiving those ideas in different ways. Those different *ways of perceiving*, Berkeley explains, are properly called “willing” and “understanding”. As it is affected by ideas, Berkeley tells us, the spirit is called “understanding”, while as it produces ideas or “operates about them”, it is called the “will”. We should note Berkeley’s insistence that these are two different terms *for the same thing*, a thinking, active principle which can perceive ideas in different ways.<sup>28</sup>

These different perceptual activities are not grounded in a metaphysical or ontological distinction (or a Cartesian modal distinction), but are rather different ways in which a single, individuated thing goes about perceiving (and, therefore, existing). Berkeley defines spirit as “that active principle of motion and change in ideas”, “that which acts”, and as something that produces certain effects (PHK §27). Berkeley explains that we have notional (non-ideational) knowledge of soul or spirit insofar as “we know or understand the meaning of those words” (PHK §27). Berkeley makes it clear that even though we cannot have an idea of spirit, that does not preclude our *knowing* it.<sup>29</sup> What does it mean to understand the meaning of the word spirit? It means understanding *what it does*. Since

26 That is, when I either perceive “external” things or else when I reflect on my ideas of memory or imagination.

27 For a discussion of the scope of our ideas’ representational content, see Bolton 1987.

28 Contrary to Locke, Berkeley does not take the mind to be a substance with distinct *powers*. EHU 2.4.2. See Roberts 2007, 103.

29 It is a mistake to think, in Berkeley’s system, having an idea of something is the only kind of knowledge. In PHK §89, he writes: “the term ‘idea’ would be improperly extended to signify every thing we know or have any notion of.” See also PHK §27, §142.



spirit is “an active principle insofar as it produces motion and change in ideas”, understanding what spirit does, in turn, requires perceiving ideas that are subject to that motion and change (PHK §138; DHP 232).<sup>30</sup> And this is exactly what Berkeley thinks we do: we frequently perceive series of ideas that could only have been caused by the will of either a finite or infinite spirit (PHK §146–7). Some of those ideas are caused by our own spirit, while some are not, and thus we learn to differentiate between what our own spirit does and does not do (§28–9).

What is it like to gain immediate knowledge of the “will” and the “understanding”? And *how* do I gain such perceptual knowledge if it is not represented by my ideas? The answer lies in experience; though we do not have ideas that represent the activity of our spirit, we are nonetheless able to experience the spirit as it acts. There is, for Berkeley, a distinction between the phenomenology of being a *willing* agent and the phenomenology of being an *understanding* one; this is dependent on the nature of the ideas perceived. In other words, self-knowledge *requires* perceiving certain ideas, but is not the same as perceiving those ideas. Rather, self-knowledge consists in perceiving those ideas in different ways (and *what that is like*).

In §§28–30 of the *Principles*, Berkeley gives a descriptive (phenomenological) account of the difference between “willing” and “understanding”. He does this by explaining what these two perceptual acts are like, as well as the kinds of ideas that are involved in these experiences. He writes:

I find I can excite ideas in my mind at pleasure, and vary and shift the scene as oft as I think fit. It is no more than willing, and straightway this or that idea arises in my fancy; and by the same power is obliterated, and makes way for another. This making or unmaking of ideas very properly denotes the mind active. Thus much is certain, and grounded on experience. (PHK §28)

In this way, Berkeley explains that in those cases where we experience the voluntary nature of our ideas – cases where we have total control over our ideas – we ought to describe the spirit as “willing” (Stoheham 2010, 497). The experience of “willing”, then, is one of being able to “excite ideas in my mind at pleasure”. Berkeley maintains that it is clear in such cases, where we can make or unmake our ideas voluntarily, that a particular perceptual activity is taking place. What’s more, Berkeley holds that it is immediately clear in such cases that something different is going on compared to when we experience sense perceptions, for the *contrast* between willing and understanding is immediately experienced. He writes:

But whatever power I may have over my own thoughts [as described above], I find the ideas actually perceived by sense have not a like dependence on my will. When in broad daylight I open my eyes, it is not in my power to choose whether I shall

---

30 For a similar claim, see Pearce 2017, 127.



see or no, or to determine what particular objects shall present themselves to my view; and so likewise as to the hearing and other senses, the ideas imprinted on them are not creatures of my will. (PHK §29)<sup>31</sup>

In this passage, Berkeley describes an instance in which the spirit is rightly called “understanding”; a case where the ideas I perceive are involuntary. I have no determination over the ideas I see, hear and so on. The experience of “understanding” is one of perceiving ideas that “are not creatures of my will”. Again, Berkeley argues that our experience of perceiving these ideas will make immediately clear the distinction between the will and the understanding. This is not the kind of knowledge that requires reflection or inference; but is immediately given. Thus, the manner in which we perceive our ideas (and what that is like) constitutes a kind of immediate data. This data does not take the form of ideas (like my data concerning the sensible world) but the form of non-ideational knowledge of *what it is like* to perceive in different ways. Knowledge via ideas is knowledge of what something (else) is like *to me*, while knowledge of the mind is knowledge of what it is like *to be me*.

It is important to note that Berkeley thinks that willing can produce more than just “ideas of imagination” in that, by means of volition, we can move our own limbs (and thereby produce a restricted set of “sensible ideas”. Beyond ideas of imagination, he explains, “the will of man has no other object than barely the motion of the limbs of his body” (PHK §147).<sup>32</sup> Bodily movements, therefore, constitute a kind of idea of sense that is voluntarily produced. So Berkeley maintains there are two distinctions in play here: (i) the distinction between voluntary and involuntary perception, and (ii) the distinction between ideas of sense and ideas of imagination. Both play some kind of role in our experience of the spirit as it is either willing or understanding. We experience *most* ideas of sense as involuntarily perceived (e.g. the sun in broad daylight), but some (e.g. the movements of my body) as voluntary. *Most* ideas of imagination are experienced as voluntarily perceived (e.g. memory, “compounding and dividing” ideas to form imaginary images; PI §10), but some are experienced as involuntary (e.g. “dreams, phrensies”; PHK §18). The latter of these cases (involuntary ideas of imagination) are the most problematic and Berkeley presumably does not have them in mind when he claims that “The ideas of sense are more strong, lively, and distinct than those of the imagination.” (PHK §30) One of these distinctions must generate our immediate awareness of the difference between “willing” and “understanding”. Since the distinction between ideas of sense and ideas of imagination does not

---

31 See also NB 499.

32 Within the literature, there remains a debate concerning “human agency” in Berkeley’s philosophy; i.e. the extent to which finite spirits can be said to act upon things in nature. The debate focuses on whether finite spirits can be said to have causal efficacy (or if Berkeley is really an occasionalist). My focus is on the way we experience “will” and “understanding” – and Berkeley is clear (in PHK §§28–9) that we *experience* ourselves as affecting some things and unable to affect others. It is not critical to my reading, then, that this debate be resolved one way or the other. For recent accounts of the debate, see (for example) Roberts 2007, 111–23; McDonough 2008.

(from the fact that ideas of sense are not exclusively involuntary), it must be the voluntary/involuntary distinction that does. After all, whether or not a possible hallucination turns out to be an idea of sense or an idea of imagination, it will not require reflection or inference to determine whether it was voluntary or involuntary. And indeed, that hallucinations are involuntarily perceived is what distinguishes them from imaginings or daydreams.<sup>33</sup>

This reading of Berkeley's account of self-knowledge as immediate experience of the ways in which we perceive ideas (voluntarily and involuntarily) has a number of strengths. Firstly, it is consistent with Berkeley's metaphysical claims, according to which there is no more to spirit than its actions. Consider, for example, the following claim from the *Notebooks*:

[the] substance of a spirit is that which acts, causes, wills, operates, or if you please (to avoid the quibble that may be made on the word *it*), to act, cause, will, operate [my emphasis]. (NB 829)<sup>34</sup>

Berkeley's point is that the substance of a spirit is not something *behind* mental acts, causes, volitions and operations, but is constituted by the *doing* of those very acts. It is for that reason that Berkeley is motivated to avoid positing a metaphysical subject (an "it") behind the actions of causing, willing, operating and so on. Recall, for example, Descartes' remark: "thought alone belongs to it [i.e. the mind]" (CSM 1:195). It is clear that Berkeley is distancing himself from such a position. For that reason, Stephen H. Daniel paraphrases Berkeley as such:

Understanding, willing, and spirit are thus not three things of which we just happen not to have ideas, for to say that would fall into the very abstractionist way of speaking about spiritual substance that Berkeley repeatedly rejects. In his eyes, just as there is no abstract material substance that underlies or supports sensible qualities, so there is no spiritual substance abstracted from willing or understanding. (2018, 661)

Daniel also treats Berkeley's account of both minds and sensible things analogously: if there no abstract "it" behind sensible qualities, why should there be an abstract "it" behind perceptual acts? In sum, there are no abstract "its" in Berkeley's philosophy at all. Consider PHK §49 where Berkeley considers the nature of a "die" and explicitly denies that "it" is something in which certain qualities (its colour, cubic shape, etc.) inhere. Rather the die

---

33 I *know* when I am voluntarily imagining things, but I don't know when I'm hallucinating until after the event – that's what is so disconcerting. Of course, if I voluntarily took a hallucinogenic I would be expecting to hallucinate (as Manuel Fasko pointed out to me). Nonetheless, (I presume) there would be some involuntariness (i.e. "reality") in the resulting experiences.

34 Similarly, see NB 658–9; 499a.

is constituted by those very properties.<sup>35</sup> In both cases, Berkeley denies there is an “it” to be known: to know a set of features (whether those be the sensible qualities of the die or the different operations of the mind) is to know the thing itself. The substance of the mind is its perceptual acts, just as the “substance” of a die is its sensible qualities.

Secondly, this reading is consistent with Berkeley’s broader claim that existence is not to be conceived of separately from perception (NB 429). Berkeley explains how the doctrine of abstract ideas has had a pernicious effect on approaches to knowledge of spirit. He argues it has given thinkers cause to believe that

they could frame abstract notions of the powers and acts of the mind, and consider them prescinded, as well as from the mind or spirit it self as from their respective objects and effects. (PHK §143)

This provides further evidence that Berkeley thinks that to know a spirit is nothing more than to know its acts. A similar sentiment is also expressed when Berkeley argues

If therefore it is impossible that any degree of those powers [of the mind] should be represented in an idea, it is evident there can be no idea of spirit. (PHK §148, 1734 edition)

The principle at work here is clearly that *to know or represent the mind’s acts is one and the same as to know or represent the mind itself*.

Finally, on this reading Berkeley’s account of self-knowledge provides us with immediate data: the individuated, immediately discernible ways in which we perceive (willing and understanding). As such, we can now make more sense of the claim that it is by virtue of self-knowledge that we come to gain mediate knowledge of other spirits (DHP 231). *By virtue of its being a spirit*, I can say of another spirit; it must, at times, be “willing” and, at others, “understanding”. Of course, I cannot necessarily make claims about *just what it is* another spirit is willing or understanding, but this is no more problematic than not being able to mediate apprehend the accidental features (e.g. colour, size) of a triangle I have not yet perceived. What’s more, I know exactly what is meant when I use these terms, even if it is in reference to another spiritual agent.

In the next section I will consider an objection to this reading of Berkeleian self-knowledge, in which two problematic claims are upheld. Firstly, that Berkeley’s account of self-knowledge is non-perceptual. Secondly, that Berkeley, in a strict sense, does not take the self to be an object of human knowledge. In particular, I will focus on Talia Mae Bettcher’s claim that Berkeleian self-knowledge is a kind of “*non-perceptual awareness*” (Bettcher 2007, 46).

---

35 For more on this, see Cummins 1963.

# Self-knowledge as Perceptual Knowledge

Thus far, I have argued that, according to Berkeley's account of self-knowledge, we can gain immediate knowledge of our own spirit by means of our immediate experience of ourselves as willing and understanding. On my reading, to be immediately aware of the manner in which we perceive ideas (voluntarily or involuntarily) is enough to constitute immediate self-knowledge. This, in turn, I have argued, makes it clearer what our mediate knowledge of other spirits is constituted by. The type of objection to this view I will anticipate in this section is likely to be derived from the following remark:

But *besides* all that endless variety of ideas or objects of knowledge, there is likewise something that knows or perceives them, and exercises divers operations, as willing, imagining, remembering about them. (PHK §2 [my emphasis])

In this passage, Berkeley seems to suggest that spirit, which “knows and perceives” ideas, is distinct from the kinds of things which can be objects of knowledge. This might suggest not only that spirits are not the kinds of things we can have ideas of, but that they are not the kinds of things we can have *any* immediate knowledge of. On this reading, it would presumably be more plausible to read Berkeley as holding that we have a notion of spirit *only* insofar as we “understand the meaning of the word” (PHK §140). That is, I can infer the existence of spirits and thereby use the term “spirit” as a placeholder for any knowledge I might have had thereof (were it possible to have such knowledge), but I can't know any more about spirit than that.

Yet, in light of the account of both self-knowledge and knowledge of other spirits outlined so far it would be very surprising if this turned out to be Berkeley's view. This much is acknowledged in Bettcher's reading, when she writes:

Does Berkeley think ‘What I am myself’ and ‘that which I denote by “I”’ is unavailable to awareness? It seems hard to believe, especially when he then writes ‘[...] we know other spirits by means of our own soul, which in that sense is the image or idea of them [...]’. (2007, 46)

So Bettcher, like me, is convinced that Berkeley must have an account of *knowledge* of oneself as spirit. But in contrast to my own presentation of Berkeley's view Bettcher then claims we can avoid this concern if we see self-knowledge as a piece of “self-evident” information; something we, in effect, get for free when we realise we are neither *nothing* nor an idea (or a collection of ideas). As she puts it, self-knowledge is a kind of “inward feeling” or, “immediate awareness that accompanies all our thinking (the ‘I’)”. In this way, Bettcher offers a reading in which the self (or spirit in general) is not strictly an object of knowledge and ought not to be understood in terms of perception. Admittedly, in some ways, Bettcher's interpretation appears to be consistent with my own. For example, she argues that if

Berkeley thinks that self-knowledge is the kind of awareness that can easily be confused for an idea, then it is most likely, “a kind of datum or inner feeling”. Likewise, she does not attribute to Berkeley an inferential account of the mind and its operations, where thought (or perception) is key to knowledge of a thinker (or perceiver) (2007, 47–8).

However, it should be clear by now that I do not accept Bettcher’s claim that Berkeleian self-knowledge is *non-perceptual*. The way that we become aware of ourselves *as ourselves*, on Bettcher’s reading, is by becoming conscious of the difference between (i) what is distinct from oneself (objects of understanding), and (ii) what is not. Bettcher places a great deal of emphasis on the *relationship* between perceivers and the perceived, and maintains that for Berkeley it is by virtue of acknowledging this relationship that we realise the self-evidence of the existence of the *relata*. For that reason, the manner of perceiving and the variability of that which is perceived (i.e. our ideas) does not play a role in providing immediate data pertaining to the self. For self-knowledge, all that is required is knowledge of the relationship between perceivers and perceived things: “We know that spirit perceives ideas and once we know that we know all that we need to know.” (Bettcher 2007, 51).

My initial concern with this reading is that since the *existence* of spirit is, for Berkeley, to be construed in terms of perception (NB 646), then there needs to be a good reason why *knowledge* of it is not. To suggest otherwise threatens to reduce self-knowledge to a kind of abstraction; an unnecessary attempt to distinguish a thing (in this case spirit) from the means by which we come to know it. Certainly, Berkeley gives us very clear reasons why one’s own spirit cannot be immediately *perceived*; because of the categorical dissimilarity between ideas (objects of perception) and spirits. But it seems to me that Bettcher takes an unnecessary step from (i) self-knowledge is not perceived, to (ii) self-knowledge is not *perceptual* (at all).<sup>36</sup> And, indeed, by forcing a distinction between the way in which a spirit exists (*percipere*) and our knowledge thereof, Bettcher is in danger of contravening Berkeley’s advice not to “abstract the existence of a spirit from its cogitation” – even if only conceptually (PHK §98). What’s more, later in the *Principles* Berkeley claims that spirits *can* be objects of knowledge: “To me it seems that ideas, spirits, and relations are all in their respective kinds the objects of human knowledge.” (PHK §89). This means that Bettcher’s interpretation of the ambiguity in §1 of the *Principles* ends up contradicting Berkeley’s more explicit claim later in that work.

Bettcher’s reading also gets us no closer to understanding how it is I use knowledge of my own spirit to gain mediate knowledge of other spirits. If self-knowledge is constituted merely by the self-evidence of the self as a subject in a perceiver-perceived thing relationship, it is hard to see how we could get any further than the inference that other spirits exist. But Berkeley wants to say more than this; his claim is that we know the *meaning* of the word “spirit” – whether applied to myself or another agent. What’s more, in the case of God’s spirit, it is key to Berkeley’s metaphysics that we know more than just that

---

36 Winkler suggests such a move might be derived from PHK §27 where Berkeley writes that spirit “cannot be of itself perceived”. See 2011, 234. As I have stressed, for Berkeley there is more to perceptual knowledge than just what is perceived; we also gain immediate knowledge of the *acts* of perception.

he exists, but that we can make claims about his nature and compare him with ourselves. I previously emphasised the importance of self-knowledge and the key role it plays in gaining mediate knowledge of both infinite and other finite spirits. As such, it ought to be seen as a deficiency in any reading of Berkeleian self-knowledge if it leaves us unclear as to what other spirits are really like.

## Conclusion

I have argued that, for Berkeley, self-knowledge is immediate knowledge of the operations of one's own spirit as it is perceiving, and that it is in our experience of willing and understanding ideas that we come to be immediately familiar with our self. The operations of willing and understanding provide us with an immediate account of the nature of our own spirit, and a model for what another spirit is like.

My concerns throughout this paper have been primarily epistemological rather than metaphysical. In light of Berkeley's claims that self-knowledge is the means by which we come to gain knowledge of other spirits, my foremost concern has been to give an exposition of Berkeleian self-*knowledge*. Nonetheless, there are some metaphysical ramifications. In particular, my reading of Berkeleian self-knowledge as non-abstract implies that (for Berkeley) the mind itself is not something abstracted from its perceptual acts. In that sense, my reading indicates that Berkeley's account of the self is non-Cartesian: there is no conceptual or ontological distinction between the self as mental substance and the acts (or operations) that substance engages in. As Stephen Daniel puts it: "for Berkeley, spirit is not *a* substance that just happens to will" (2018, 664). But that need not render Berkeley's account of spirit Humean or Spinozistic either: knowledge of oneself is *not* reducible to the "bundle" of ideas that one perceives.<sup>37</sup> Self-knowledge, I have argued, consists in identifying the different ways that perception occurs and *what that is like*. My reading thus implies that the self is a set of perceptual *acts*, each of which has distinct phenomenological content. That phenomenological content is dependent on, but *not* reducible to, the ideas perceived. For Berkeley, experience is not only key to knowledge but a kind of knowledge in itself. Hence the experience of *perceiving* is also a kind of knowledge. It is for this reason that Berkeley claims that the objects of human knowledge are derived not only from sensation, imagination and memory but also by "attending to the passions and operations of the mind" (PHK §1).

When Berkeley claims we know the "meaning of the word" "spirit" or "mind", he does not mean that the terms signify distinct ideas nor any kind of "abstracted" or absolute knowledge (PHK §142; AMP 7.13). Rather, his view is that I know what the word "spirit" means because I know what my own spirit is like – I know what it is like to be *me*. Sometimes it voluntarily perceives either ideas of the imagination or, in a much more limited

---

<sup>37</sup> See Daniel 2018, 668 for a similar claim.

way, sensible ideas (those that make up the collection of ideas which constitute my body). When my spirit perceives in this way, it is called the “will”. At other times, my spirit involuntarily perceives ideas, such as the sun in broad daylight, my desk when I enter my study, or the sounds another human agent makes when they talk to me. In such cases, my spirit is called the “understanding”.

In this way, when I make the claim that another spirit exists (whether divine or human) I am doing more than just making an ontological claim, I am saying that it exists in a certain way. Once I have good evidence to believe that other spirits exist and that they are indeed spirits,<sup>38</sup> I am able to make the further claim that they *will* and *understand*. This reading therefore allows for a more robust understanding of Berkeley’s claim that we have a “notion” of spirit, “in as much as we know or understand the meaning of those words” (PHK §27).<sup>39</sup>

## Bibliography

### Primary Sources

- NB Berkeley, George. 1989. *Philosophical Commentaries*. Edited by George Hasson Thomas. Ohio: Garland.
- MI Berkeley, George. 1987. *George Berkeley’s Manuscript Introduction: An Editio Diplomatica*. Edited by Bertil Belfrage. Oxford: Doxa.
- PHK Berkeley, George. 2008 [1710]. “A Treatise Concerning the Principles of Human Knowledge.” In *Philosophical Writings*, by George Berkeley, Cambridge Texts in the History of Philosophy, edited by Desmond M. Clarke, 67–149. Cambridge: Cambridge University Press.
- PI Introduction to PHK.
- DHP Berkeley, George. 2008 [1713]. “Three Dialogues Between Hylas and Philonous.” In *Philosophical Writings*, by George Berkeley, Cambridge Texts in the History of Philosophy, edited by Desmond M. Clarke, 151–242. Cambridge: Cambridge University Press.
- AMP Berkeley, George. 2010. “Alciphron, or The Minute Philosopher.” *Berkeley’s Alciphron: English text and essays in interpretation*, edited by Laurent Jaffro, Genevieve Brykman and Claire Schwartz. Hildesheim: Georg Olms Verlag.
- CSM Cottingham, John, Stoothoff, Robert, Murdoch, Dugald and Kenny, Anthony (trans.). 1984–1991. *The Philosophical Writings of Descartes*, 3 vols. Cambridge: Cambridge University Press. [Citations are to volume followed by page number in that edition.]

---

38 For discussion of Berkeley’s answer to the “problem of other minds” and how we know they exist, see (for instance) Bennett 2001; Falkenstein 1990; Fogelin 2001.

39 Thanks to Kenneth Pearce, Manuel Fasko and two reviewers from this volume for comments on earlier drafts of this paper.



## Secondary Sources

- Adams, Robert. 1973. "Berkeley's 'Notion' of Spiritual Substance." *Archiv für Geschichte der Philosophie* 55: 47–69.
- Bennett, Jonathan. 2001. *Learning from Six Philosophers: Volume 2*. Oxford University Press: Oxford.
- Bettcher, Talia Mae. 2007. *Berkeley's Philosophy of Spirit: Consciousness, Ontology and the Elusive Subject*. London: Continuum.
- Bolton, Martha Brandt. 1987. "Berkeley's Objection to Abstract Ideas and Unconceived Objects." In *Essays on the Philosophy of George Berkeley*, edited by Ernest Sosa, 61–81. Boston: Springer.
- Cummins, Phillip D. 1963. "Perceptual Relativity and Ideas in the Mind." *Philosophy and Phenomenological Research* 24(2): 202–14.
- Daniel, Stephen H. 2008. "Berkeley's Stoic Notion of Spiritual Substance." In *New Interpretations of Berkeley's Thought*, edited by Stephen H. Daniel, 202–33. New York: Prometheus Books.
- Daniel, Stephen H. 2018. "Berkeley's Non-Cartesian Notion of Spiritual Substance." *Journal of the History of Philosophy* 56(4): 659–82.
- Falkenstein, Lorne. 1990. "Berkeley's Argument for Other Minds." *History of Philosophy Quarterly* 7(4): 431–40.
- Flage, Daniel. 1987. *Berkeley's Doctrine of Notions: A Reconstruction based on His Theory of Meaning*. London and New York: Croom Helm and St. Martin's Press.
- Flage, Daniel. 2008. "Berkeley's Epistemic Ontology: The *Three Dialogues*." In *New Interpretations of Berkeley's Thought*, edited by Stephen H. Daniel, 45–76. New York: Prometheus Books.
- Fogelin, Robert J. 2001. *Routledge Philosophy Guidebook to Berkeley and the Principles of Human Knowledge*. Routledge: London.
- McDonough, Jeffrey K. 2008. "Berkeley, Human Agency, and Divine Concurrentism." *Journal of the History of Philosophy* 46(4): 567–90.
- Holden, Matthew. 2017. "Berkeley on Inconceivability and Impossibility." *Philosophy and Phenomenological Research* 98(1): 107–22.
- Pearce, Kenneth L. 2017. *Language and the Structure of Berkeley's World*. Oxford: Oxford University Press.
- Roberts, John Russell. 2007. *A Metaphysics for the Mob*. Oxford: Oxford University Press.
- Russell, Bertrand. 2001 [1912]. *The Problems of Philosophy*. Oxford: Oxford University Press.
- Thiel, Udo. 2011. *The Early Modern Subject*. Oxford: Oxford University Press.
- Winkler, Kenneth. 2011. "Marvellous Emptiness: Berkeley on Consciousness and Self-Consciousness." In *Berkeley's Lasting Legacy: 300 Years Later*, edited by Timo Airaksinen and Bertil Belfrage, 223–50. Cambridge: Cambridge Scholarly Publishing.





# NATURAL LAW



# The Space for Emotions in Natural Law

## Introduction: Emotions as *passiones animae*

Space for emotions? Even if one takes natural law in a broad sense, i.e. that natural law includes both *ius gentium* / *law of nations* and *common law*, it seems more or less impossible to find a space for a concept of emotionality at first glance. The task for justice is formal rational communication in a society. Has that anything in common with emotions?

Moreover, concerning emotions in traditional natural law, one faces severe terminological problems: The term *emotion*, used in its modern sense, has no Latin equivalent, even though it sounds quite Latin. In Latin, *emotio* means *uprising* and *rebellion*. The Latin equivalents of the modern “emotions” are *affectiones* or *passiones animae*. Both, *affectio* and *passio*, in origin mean passivity, i.e. the suffering of the soul. Even though passions and virtues share a common field of meaning, passions are not emotions in the precise sense of the word. Emotions are active, they move and shatter the soul, they pretend to be spontaneous and have their own rights.

Is there any space for such a spontaneity in natural or common law? In order to find such a space, do we have to turn to affections and passions, e.g. to fear and anxious mistrust? These passions, indeed, would fit Hobbes’s concept of the individual’s natural rights, but they are strictly contrary to the traditional scholastic concept of natural law and justice.

Considering these difficulties, it seems appropriate to approach the subject historically and to distinguish three stages in the development of natural law.

1. Natural law and the space of affections according to Thomas Aquinas’s theories.
2. Thomas Hobbes’s critique of the scholastic concept of natural law and the new definition of affections in his political theory.

3. Samuel Pufendorf's defence of the traditional Jus Naturae and his creation of a space for civil activities and virtues independent from impacts of State and church.

## Thomas Aquinas Defining Natural Law

Aquinas is the most important and most influential theoretician of natural law in the Middle Ages; and in the Early Modern Period he is still eminently important. In his *Summa Theologiae*, Ia IIae qu. 90–100, he discusses the problems of natural law in detail.

### Ius sive iustum

The term *Ius* may be rendered in the English language as *law* or *justice* – neither translation meets the term *Ius* precisely.

Aquinas defines *Ius – sive iustum*, as “aliquid opus adaequatum alteri secundum aequalitatis modum”.<sup>1</sup> The pivotal term here is equality. “Equality” here has a wide range of meanings: e.g. corresponding to someone else's substance or habit, it includes proportionality, similarities, or a comparable value, it means commensurable, righteous, honest. Notoriously it turns out: in practical philosophy, the terms are not precise, they are weak and somehow cloudy – but exactly this terminology fits practical reason.

The definition of *ius positivum* is self-evident: This law is given by a lawmaker and receives its validity from the authority and might of this lawmaker.

Aquinas distinguishes *ius positivum*, *ius divinum* (divine law) and *ius civile* (civil law).

*Divine law* is the positive law as revealed in the Bible; it coincides partly with natural law.

*Ius civile* is the positive secular law; here Thomas discerns *ordinatio principis* and *conductum populi*.

The *ordinatio principis* depends on the will of the prince and receives its legitimacy from his might, his authority and his legitimate office.

*Condictum Populi* is the second source of the legitimacy of positive law. “*Condictum*” can be approximately translated as “agreement”. It coincides with *ius gentium* “cum totus populus consentit”. Thomas has two natural and reasonable arguments for the legitimacy of the *conductum populi*: 1. “quia homo naturaliter est animal sociale” (Aristotle) and 2. *ius gentium* aims at the “bonum individuale et commune”.

---

1 Thomas, S. Th. 2.2.qu.57. art. 2 in corp. *ius, sive iustum*, est aliquid opus adaequatum alteri secundum aequalitatis modum: Dupliciter autem potest alicui homini esse aliquid adaequatum: Uno quidem modo ex ipsa natura rei, puta cum *aliquis tantum dat, ut tantumdem recipiat*, & hoc vocatur ius naturale.

## Ius naturale

Thomas avoids, it seems to me, the formula “*suum cuique tribuere*” as the key formula for natural law – instead he defines *adaequatio*, equality, as a form of a contract: *cum aliquis tantum dat, ut tantumdem recipiat*.

The central chapter concerning the definition of natural law is Ia IIae Qu. 94 Art. 2 in the *Summa*: *Utrum lex naturalis contineat plura praecepta vel unum tantum*.

Thomas explains: Even though there are many particular precepts in natural law, they all can be reduced to the key principle that every being strives for the good and abhors the evil.

To prove this thesis Thomas proceeds from the Aristotelian principle that every action tends to a formal good; and because of such a formal good every action comes to pass. In Aquinas’s words: “*ita bonum est primum quod cadit in apprehensione practicae rationis, quae ordinetur ad opus. Omne enim agens agit propter finem, qui habet rationem boni.*” The consequence is: *Bonum est quod omnia appetunt*.

Exactly this principle is the core of natural law.

Aquinas unfolds his concept of natural law in the following pivotal points:

1. All living beings have an inclination (*inclinatio*) to follow their natural good. Human beings have in common with other living beings that they strive for *conservationem sui esse secundum suam naturam* (conservation of their being). This is a basic natural law and principle of every life, including human life.<sup>2</sup>

2. Every human being has, insofar as it is a sexual being, an inclination to the connection between male and female, and this inclination entails the education of their children. With these principles Aquinas states that the family is the natural basis of every society. Thomas follows here Isidor of Sevilla’s *Etymologiae*, who, in his definition of natural law, emphasizes the role of the family.

3. Rationality has a theoretical and a practical branch. Every rational being – which includes every human – strives first for cognition of the divine truth, and second for knowledge about his or her life in a human society. The human faculty of practical rationality becomes actual in the virtues (Ia IIae Qu. 94 Art. Art. 3).

For Aquinas this principle includes that man hates ignorance and does not offend others, since he has to communicate with them (Ia IIae Qu. 94 Art. 2).

4. Natural law is stable in its primary principles. These may be formulated in different ways, but must remain faithful to the spirit of the first principles (Ia IIae Qu. 94 Art. 5). Therefore, the first principle of natural law is identical in every human society, the secondary principles, however, vary (Ia IIae Qu. 94 Art. 4). These variations of the secondary principles feature particularly in common law and *ius gentium*.

---

<sup>2</sup> Ia IIae Qu. 94 Art. 2 *Secundum hanc inclinationem [ad suum esse conservandam] pertinent ad legem naturalem, ea per quae vita hominis conservatur, et contrarium impeditur.*

5. The principles of natural law are anchored in the hearts of every man; variations are possible only in minor details (Ia IIae Qu. 94 Art. 5).

## Passions and virtues<sup>3</sup>

### Justitia, Temperantia, Fortitudo, Prudentia

Practical rationality becomes actual in the virtues and has its place in social practice. Men and women have virtues (cf. S. Th. Ia IIae qu. 94, art. 3) because they are social animals; therefore, the virtues as the practice of natural law play an important role in civil law. Traditionally, four “pivotal” virtues are the most important: *Prudentia*, *Iustitia*, *Temperantia*, *Fortitudo*.

Among these pivotal virtues *justice* (*iustitia*) as *iustitia legalis* is peculiarly important, since it expands from the individual and ethical sphere into the political one.

The other virtues also play their part in the *ius gentium* – insofar as human ethics (or rather common habits) are part of the common law or law of nations/*ius gentium*.

Virtues are spiritual motive powers and in order to make their potentiality real they have to overcome hindrances, namely the passions. Passions in this sense are blind and wild emotions; virtues act to tame emotions and this task is the duty of practical rationality. To the extent that passions or emotions play a negative role in practical philosophy, they oppose the rational virtues, since they are by definition irrational. This is the place for *temperantia* taming *concupiscentia*, and *fortitudo* taming *irascibilitas*.

*Prudentia* is the virtue that by definition tends towards the good; it rationally directs the free will which can tend to evil as well as to the good.

---

3 In the *Summa Theologiae* he writes (1.2, qu. 61, art. 2 in corp.): “Principium enim formale virtutis, de qua nunc loquimur, est rationis bonum; quod quidem dupliciter potest considerari: uno modo secundum quod *in ipsa consideratione rationis* consistit; & sic erit una virtus principalis, quae dicitur *prudentia*; alio modo, secundum quod circa aliquid ponitur rationis ordo: & hoc vel circa operationes; et sic est *Justitia*: vel circa passiones; & sic necesse esse duas virtutes: *ordinem enim rationis necesse est ponere circa passiones*, considerata repugnantia ipsarum ad rationem: Quae quidem potest dici dupliciter: uno modo secundum quod passio impellit ad aliquid contrarium rationi; & sic necesse est quod passio reprimatur, & ad hoc denominatur *temperantia*; alio modo secundum quod passio retrahit ab eo quod ratio dictat, sicut *timor* periculorum, vel laborum: & sic necesse est quod homo firmetur in eo quod est rationis ne recedat, & ad hoc denominatur *fortitudo*.

“Et similiter secundum subiecta idem numerus invenitur: Quadruplex enim invenitur subiectum huius virtutis, de qua nunc loquimur, scilicet rationem per essentiam, quod prudentia perficit; & ratione per participationem, quod dividitur in tria: id est voluntatem, quae est subiectum iustitiae: & in *concupiscibilem*, quae est subiectum temperantiae: & in *irascibilem* quae est subiectum fortitudinis.”

## Charity

Faith, hope and charity are, according to St Paul, the three spiritual virtues. Among these virtues only charity can be regarded as emotional; however, as far as I see, Aquinas does not include charity in the field of natural or common law. Charity, however, plays a particular role in the later tradition of natural law. Francesco Suarez, later, emphasizes that Charity, as love of God, is like a spiritual pious emotion. Leibniz uses charity in his most important definition of justice: *Justitia est caritas sapientis*. (Justice is the charity of the wise). And in Pufendorf's political theory, Charity will appear in the guise of the key concept of *sociabilitas*.

## Ludi – games

Of course, natural law and *ius gentium* have ethical and sociological dimensions. Especially in the field of *ius gentium*, practical social habits play an important role. Here it seems worthwhile to mention that ethics entails practical habits, which can be subsumed under the term *sociabilitas*, although Aquinas, to my knowledge, does not use this term. However, as an Aristotelian, Thomas acknowledges how important institutions of pause, leisure and relaxation are in order to fulfil the duties towards God and man.

One finds Aquinas's thoughts in his considerations on *Ludi* – playing and games. Thomas considers games as necessary exercises of natural virtues, insofar as they are helpful against afflictions of vices or other affects.<sup>4</sup> He proceeds from the experience that after hard work mind and body become tired and need recreation. He quotes Aristotle's Nicomachian Ethics 4 that we need in the course of our daily life some pause with games (“quòd in hujus vitae conversatione quaedam requies cum ludo habentur”). Playing games should not lead to laziness and should remain within the framework of prudential reason. If such boundaries are respected, playing games could become a virtue. “Habitus autem secundum rationem operans est virtus moralis. Et ideo circa ludos potest esse aliqua virtus.” (S. Th. 2.2. qu. 168 art. 2 corp. 4 sent. D. 16. Qu. 4 art. 2 qu.1 in corp.). Thomas proceeds even further; he recommends having fun with the games when he quotes Cicero's rhetorical advice:

Fun has to be congruent with work and persons; therefore, Cicero says (I Rhet): If the audience tire it is useful to start with a new subject or with a joke, insofar as the dignity of the subject allows it.<sup>5</sup> (“iocosa debent congruere negotiis & personis: unde & Tullius dicit in I Rhet. Quòd quando auditores sunt defatigati non est inutile aliqua de re nova, aut ridicula oratorem incipere, si tamen rei dignitas non adimit jocundi facultatem.”) (S. Th. 2.2. qu. 168 art. 3 ad 1)

---

4 *Moralis Divi Thomae [...] Ex omnibus ipsius Operibus deprompta. Autore Ludovico Bancel. Tom. 2, Venedig 1723, 66.*

5 All translations are mine, unless marked otherwise.



# Thomas Hobbes: Monopolizing Positive Law

## The abandoning of the traditional Natural Law

Thomas Hobbes attacks the traditional concept of natural law severely. He no longer accepts any idea of a natural law, which obligates mankind in order to direct the societies to their public and individual good. He especially does not accept Aristotle's guiding principle that men are by nature peaceful political beings. For Hobbes, all laws whatsoever receive their legitimacy only from the sovereign to whom, in the contract of submission, everyone gives up all his or her individual rights. No public good exists any longer. Laws are enacted in order to foster the stability of a state, and the task of the state is to fence in the individuals' war against each other. Mistrust and fear are the emotions that force mankind to undergo this contract of subjugation under an absolute ruler.

Therefore, Hobbes turns all traditional concepts of Natural Law upside down.

True Reason is a certaine Law, which (since it is no lesse a part of Human nature, than any other faculty, or affection of the mind) is also termed natural. Therefore the Law of Nature, that I may define it, is the Dictate of right Reason, conversant about those things which are either to be done, or omitted for the constant preservation of Life, and Members, as much as in us lye. 2. But the first and fundamental Law of Nature is, that Peace is to be sought after where it may be found; and when not, there to provide ourselves for helps of war. (Hobbes 1651; *De Cive*, *Preface to the Reader*)

That is, by the way, a parody of the first chapter of Grotius's "De iure belli ac pacis", namely that a war is begun in order to attain peace ("bellum pacis causa suscipitur").

In a footnote in the *Preface to the Reader* of *De Cive* Hobbes explains his concept of the natural state.

By Right Reason in the natural state of men, I understand not, as many doe, an infallible faculty, but the act of reasoning, that is, the peculiar and true ratiocination of every man concerning those actions of his which may either redout to the damage, ore benefit of his neighbours. (Hobbes 1651; *De Cive*, Def. law II, 1.)

Here the pessimistic mirroring of Aquinas's optimistic conception of the *status naturalis* is obvious, after which every action tends by its nature to an individual or common good. Hobbes reduces the term "good" to "benefit" and its contrary, "damage". Herewith he abandons every moral dimension in the concept of law; law is just a pragmatic means to suppress human natural behaviour.

## **Fear and mistrust are the affections of the natural state of the human race, not consociatio**

Hobbes's concept of *status naturalis* is defined as the state outside and before the constitution of civil society. Civil society for Hobbes always means the absolute state. Against the background of a civil contract of subjugation under a sovereign authority, fear and mistrust determine the natural human state. Hobbes's new interpretation of the *status naturalis* replaces Aquinas's traditional concept of natural law, which pretended to be valid for mankind without exception. Hobbes's oppositional theory claims to replace Thomas's concept and to be as valid and universal as the now bygone scholastic one.

In his preface of *De Cive* Hobbes writes:

I set down for a principle by experience known to all men, and denied by none, to wit, that disposition of men are naturally such, that except they be restrained through fear of some coercive power, every man will distrust and dread each other, and as by natural right he may, so by necessity he will be forced to make use of the strength he hath, toward the preservation of himself. (Hobbes 1651; *De Cive, Preface to the Reader*)

Therefore, the *status naturalis* consists in the famous "bellum omnium contra omnes". Hobbes shows

that the state of men without civill society (which state we may properly call the state of nature) is nothing else but a mere warre of all against all; and that in that warre all men have the right unto all things. (Hobbes 1651; *De Cive, Preface to the Reader*)

One decisive element of Hobbes's political theory is the extreme individualism of every human being; there is no civil binding whatsoever in the *status naturalis*, neither charity nor benevolence nor coherence of the family. This social atomism results in the only right that is valid in the state of nature, namely, that *all men have a right unto all things*: This phrase defines the new, Hobbesian natural law as individualized power without any moral restrictions.

## **Free will tends to gain and glory, benevolence ceases to be a political category**

For Hobbes it is obvious; every society exists either for gain or for glory of an individual human being and certainly not for the love of our fellow beings, but only for love of ourselves. He asserts: Free will tends to the pleasure of the senses or to the pleasure of the mind. The

pleasure of the senses consists in conveniences, the pleasures of the mind in honour. There is no space for mutual benevolence, only for fear and terror.

Therefore, affects and emotions reign over every society unless they are tempered by the positive law of the sovereign. Every independent natural law is abandoned. Hobbes explains this train of thought as follows:

The cause of mutuall fear consists partly in the naturall equality of men, partly in their mutuall will of hurting: whence it comes to passe that we can neither expect from others, nor promise our selves the last security. (Hobbes 1651, I, 3)

So the only consequence is to subjugate the individual freedom under the rule of an absolute power which, in positive laws, defines and protects every individual freedom – nothing whatsoever remains for traditional natural law.

## Samuel Pufendorf: Sociabilitas as Natural Habit of Civil Life

Pufendorf's main attempts are to define law as both a rational and natural institution. In order to achieve this aim he defends the traditional law of nature and the concepts of common law against Hobbes's attacks. In the title of his chief work, *De iure Naturae et gentium libri octo*, (1672) his intellectual claim is obvious. He wants to prove the universal validity of natural law, to adapt the traditional topoi of law to the constitutions of the monarchies in the seventeenth century and to establish a realm of civic life that has its own rules. For this reason, he develops a new theory of sociability independent from churchly and political impacts.

### Entia moralia

For Pufendorf as for Aquinas, human nature is essentially a rational one, and the practical version of this rationality is moral consciousness. Pufendorf explains this train of thought by his leading concept of *entia moralia*. *Entia moralia* are practical ideas of our moral consciousness which moderate the wild affects of our soul and coin the habit of moral behaviour. Pufendorf is therefore convinced that human consciousness entails a general knowledge of natural rights and obligations. Those *entia moralia* are a general mark of the free actions of all men, and one could therefore recognize them throughout the course of history.

Pufendorf conceives of *entia moralia* as regulative ideas of law that result from historical moral and legal experiences. To press it into a formula: *entia moralia* are central topoi of right and law. Pufendorf defines them as

certain modes, superimposed by intelligent beings unto outside physical experiences in order to tame and temperate free and arbitrary human actions and render them into a human life of order and decorum. (“Entia moralia sunt „modi quidem, rebus aut motibus physicis superadditi ab entibus intelligibilibus, ad dirigendam potissimum et temperandam libertatem actuum hominis voluntariorum, & ad ordinem aliquam ac decorem vitae humanae conciliandum”). (Pufendorf 1998, I, § 3, 14)

This function of the *entia moralia* presupposes that men are moved and driven by affects, and that these affects are the living forces that constitute the might of free human acts. In order to harmonize this unrestrained freedom of the corporeal driving forces with the duties towards God, neighbour and oneself, they have to be appeased. Pufendorf states, therefore, right at the beginning of *De Iure Naturae et Gentium* as an axiom of natural philosophy (not far from Hobbes, by the way), that by nature all beings have

affections from which derive the infinite varieties of emotions which we experience in this our universe. (Affectiones, “per quas producuntur infinitae illae motuum varietates, quibus omnia in hocce universo agitari cernimus”). (Pufendorf 1998, I, I, § 2, 14)

These natural affects are the forces of life which, by the *entia moralia*, are directed towards the commodities of civil life.

## Pufendorf against Hobbes

Pufendorf looks upon his major work *De Iure Naturae et Gentium* as being largely directed against Hobbes’s destruction of the traditional natural law. He states that it does not correspond to human nature to live outside the law (Pufendorf 1998, “exlex” Book 2, chapter 1). This is one implication of Pufendorf’s concept of *entia moralia*. His general principle of natural law is

that every man substantially is due to cultivate and preserve peacefulness and *socialitas* against the other, which is the essence and aim of every human community in the world. (“cuilibet homini, quantum in se, colendam & conservandam esse pacificam adversos alios socialitatem indoli & scopo generis humani in universum congruentem.”) (Pufendorf 1998, § 17, 145)

The concept of *socialitas* is Pufendorf’s central argument against Hobbes; it coincides with the common European Aristotelian idea of man as *zoon politicon*. Pufendorf supports this concept with a theological argument:

In order to supply the reasonable dictations with the force of law, it is necessary to presuppose that God exists, that He has universal knowledge and governs the human race (“ut ista rationis dictamina obtineant vim legum, necessum sit supponere DEUM esse, & per ipsius providentiam tum omnia, tum imprimis genus humanum gubernari.”) (Pufendorf 1998, § 19, 153)

The decisive argument is that God’s authority corroborates the natural law of socialitas and that it threatens disobedience to it with sanctions. This argument is directed as much against Hobbes, who had abandoned all natural duties and had transferred all authority to the absolute ruler, as against Grotius, who had stated that natural law is valid “etsi Deus non daretur”.<sup>6</sup> For Pufendorf, obligation is the main concept in natural law:

The obligation of natural law stems from GOD. In order that the dictamina rationis receive legal power, it is necessary to presuppose a sublime principle. (“Obligatio legis naturalis est a DEO. Enimvero ut isthac rationis dictamina vim legis obtineant, sublimiore principio opus est.”) (Pufendorf 1998, § 20, 154)

For Pufendorf, only a transcendent will could grant the obliging *vinculum*.

## The doctrine of duties

Pufendorf identifies three fields of duties: duty towards God, towards the neighbour and towards oneself (cf. Keriting 1989).

### The duties towards God

The duties towards God, which Pufendorf describes in his booklet *De officio hominis et civis*, a popular abridged version of *De Jure Naturae et Gentium*, corresponded to the theses which Herbert of Cherbury had formulated as ethical axioms of Deism.<sup>7</sup> Pufendorf accepted them completely; however, he did not take over Herbert’s concepts of transcendent rewards and punishments. Pufendorf enlarged Herbert’s axioms with some more precise predications of the divine: So he defined six duties towards God, four theoretical and two practical ones (*De officio hominis et civis* 1,4).

The theoretical articles: 1. God exists. 2. God is the creator of the world. 3. God governs the whole world and peculiarly the human race. 4. God has no imperfect attribute. The

---

6 Pierre Bayle’s provocative idea of a pious and virtuous state of atheists is not Pufendorf’s option. Bayle: *Réponse aux questions d’un provinciale*. 5 vols. Rotterdam: Leers 1704–1707.

7 “1. Esse supremum aliquod Numen 2. Supremum istud numen debere coli. 3. Virtutem cum pietate coniunctam [...] praecipuam partem Cultus Divini habitam esse et semper fuisse. 4. Horrorem scelerum Hominum animis semper insedissee; Adeoque illos non latuisse Vitia et scelera quaecunque debere ex poenitentia. 5. Esse praemium, vel poenam post hanc vitam.” (Herbert of Cherbury 1645, 210–22).

practical duties consist in internal and external worship. 5. In his personal inner worship man has to give God due honour, and 6. external worship consists in obeying and praising God.

This is a very broad and mild definition of the duties towards God. For Pufendorf as well as for the whole traditional law, every duty corresponds to a right. Consequently, the duties towards God have as their pendant the right of worship. For Pufendorf the veneration of God is therefore a natural human right; however, he does not pretend that natural law prescribes a peculiar form of veneration. This is the basis of Pufendorf's concept of tolerance.

## Duties towards oneself: the sphere of gentle affects

Pufendorf divides the duties towards oneself into those towards the soul and those towards the body. The duties towards the body encompass striving for a comfortable living standard, health and the right of self-defence.

He lists three duties towards the soul: 1. the pious edifying exercise of good will, i.e. the free acts for good practice and virtue. 2. Striving for outward goods and property, insofar as this property is compatible with the general aims of the community, the *bonum commune*. 3. Good civil and bourgeois life harmonizing with the civil habits. This includes, as Pufendorf points out, amusements and entertainment ("Belustigung und Ergötzlichkeit").<sup>8</sup> Here is the space for gentle affects.

Partaking in common amusements, erudition and *studia humanitatis* is part of the duties towards oneself. This is the sphere of *sociabilitas*, in which Pufendorf is peculiarly interested. Pufendorf emphasizes in his booklet *De officio hominis et civis* as well as in *De Jure Naturae et Gentium*

that man is an 'animal sociabile' – a social being – because only he, different from other animals, is able to arrange a comfortable life for himself, and that there exists no other being which has that ability. ("dass der Mensch ein Gemeinshaftswesen (animal sociabile) sei, weil er, anders als andere Lebewesen, fähig sei, das Leben confortabel zu gestalten, und daß es sonst kein Wesen gebe, das dazu ebenfalls in der Lage sei.")<sup>9</sup>

---

8 1,5, § 6 In the German translation of "De officio hominis et civis" of Immanuel Weber, which Pufendorf supervised. This translation is also printed in Gerald Hartungs' edition of *De officio*, vol 2 of Pufendorf's *Gesammelte Werke*, Berlin 1997.

9 Cf. JNG II, 3, § 15: "Nam etiam ideo animal sociabile hominem diximus, quod homines mutua commoda, magis quam ullum animantium, promovere idonei sunt; sicuti & contra nullum animal plus commodi ab homine experiri potest, quam ipse homo."

## The three duties towards the neighbour

*The three duties towards the neighbour* quote Ulpian's principles of Roman law:

Neminem laedere, suum cuique tribuere, honeste vivere.

Especially the first duty, "Nobody should hurt his neighbour", is immediately directed against Hobbes's doctrine of aggressive affects. The second duty consists in "that everybody should hold the other as equal ("daß jeder den anderen für seines gleichen halten solle"). (De officio I, cap, 7: De agnoscenda naturali hominum aequalitate) With this principle Pufendorf emphasizes the equality of men in a civil society as a duty, and this principle implies the original equality of man as a natural human right. The third duty is again the duty of *consocialitas* – sociability. I shall come back to it.

If one regards this programme of human duties and rights carefully, Pufendorf's central idea is obvious, namely that human morality is characterized by the knowledge and practice of *entia moralia*, i.e. of duties and rights that are based in human consciousness and experience. This morality tempers the natural affections; and as the result of this process, a civil society becomes desirable as a sphere of comfort and virtue. That is what Pufendorf has in mind when he praises the concept of "*persona moralis*".

## The civil sphere of *cultura animi*

Pufendorf described his ideal of civic culture in 1672 in a chapter of the first edition of *De Jure Naturae et Gentium*, and he enlarged that chapter in the second edition of 1684. I quote this second edition.

Chapter 4 in Book 2 has the title: On the duties of man towards himself, concerning the *cultura animi* as well as the care of body and life. ("De praestantibus hominis adversus seipsum tam circa culturam animi, quam curam corporis et vitae.") (DJNG II, 4, 163–185) Here, Pufendorf describes the sphere of *sociabilitas* in detail; and he emphasizes that *sociabilitas* forms a rather independent space of civil society within a constitutive monarchy. Pufendorf's idea of this civil space is not quite the same as Rousseau's idea of a "*société civile*" or Hegel's concept of *Sittlichkeit*. However, with his concept of *sociabilitas*, Pufendorf lays the groundwork for the theories of his followers.

The concepts of the *entia moralia* structure the sphere of *cultura animi*; the *entia moralia* feature as habitual obligations, which receive their authority by the divine will. Therefore, the sphere of *cultura animi* is first characterized by natural theology, and Pufendorf has the function of religion in mind, which grants the natural law. He writes:

Among the sentences which are to be sucked in by the human race, the central principle is that God created and still rules this universe." ("Inter sententias igitur, quas penitus imbibisse omnes homines oportet, princeps est ea, quae circa eum tamquam Creatorem et rectorem huius universi versatur.") (DJNG II, 4, §3, 164)

The acceptance of this principle is “*praecipium hominis officium*” (DJNG II, 4, §3, 164); atheism and Epicureanism can therefore not be tolerated. Consequently, the veneration of God is a human duty. It seems that Pufendorf has no particular state religion or denomination in mind, and in *De Jure Naturae et Gentium* he only uses the concept of Deism. However, Pufendorf does not tolerate Catholicism, because the Catholic denomination does not accept the sovereignty of the state in questions of cult and worship.

The culture of the mind (*cultura animi*) and a certain ambition in civil society are duties of every citizen. Pufendorf knows well that a society needs the possibility of advancement; here, of course, the framework of decorum has to be respected. Pufendorf writes:

Within a culture it is a most important task, to know how to put a just price on the things which humans seem to aim at. This means that one knows what one can with decency strive for. Among these aims, those are appreciated as most desirable that are apt to heighten the social reputation, i.e. the opinion of distinction and excellence, from which honour and glory derive. (“*Est quoque illa culturae pars maxime necessaria, ut quis justum pretium rebus, quae adpetitum humanum praecipue simulant, ponere norit. Ex hoc quippe dependet, quantum circa quaque earum conniti deceat. Inter eas igitur judicatur vel splendissima, & quae erectioris hominis indolis praecipue sollicitare idonea est, opinio praestantiae & excellentiae, unde honor & Gloria progignitur.*”). (DJNG II, 4, §9, S. 170 f)

It goes without saying that these civil ambitions have to respect the rules of natural law and follow the advice of the *entia moralia*.

Here, too, one finds the space for “*voluptates*”, insofar as they do not contradict the natural obligations.<sup>10</sup> It is the virtues that can be characterized well as serene and habitual. They temperate the affects and turn them into good habits. Therefore, it is obvious for Pufendorf that happiness and serenity perfectly fit human nature (“*Laetitia in se naturae cumprimis congruit.*”) (DJNG II, 4, §12, S. 174); and it is, of course, also natural to restrain sadness.

Love is a friendly affect of human nature, but one has to temper it by means of reason, so that it is appropriate to its object. (“*Amor naturae humanae amicus affectus, sed qui ratione hactenus est temperandus, ut in obiectum dignum feratur.*”) (DJNG II, 4, §3, S. 164)

The affects of hate and envy are to be suppressed by education; and one should moderate the affect of hope so that it does not become “*vana, incerta aut supra vires posita*” (DJNG II, 4, §3 S. 164). Fear, quite differently from Hobbes, is characterized as an enemy of the human mind and as a completely unnecessary emotion, “*humanae mentis inimicus,*

---

10 Pufendorf deals with the virtues in detail only in the second edition of DJNG; perhaps the success of *De Officiis* encouraged him to enlarge this chapter.



planeque inutilis” (DJNG II, 4, §3, S. 164). Also wrath (*ira*), is described as “*violentissimus juxta ac perniciosissimus affectus*” (DJNG II, 4, §3, S. 164).

When he recommends *Studia literaria* as part of civil society, it looks as if Pufendorf is recalling his youth, when he studied extensively for a long time – more than ten years – at the universities of Leipzig and Jena, where he joined the *societates literales* reading classical literature (DJNG II, 4, §13, 175).

The sphere of *cultura animi* is perhaps Pufendorf’s most important contribution to modern political theory, because it already shapes the field of civil freedoms within the state, be it a constitutional monarchy or a bourgeois republic. However, this political progress – if one believes in progress – is due to very traditional ideas of natural rights, namely the right of playing and gaming. And here recalling Thomas Aquinas is helpful. For him the right of relaxation, playing games and the serenity of mind was a substantial aspect of human nature and it was – and is – consequently a part of human rights.

Perhaps such habits are the most important ones for a good and happy life.

## Bibliography

### Primary Sources

Aquinas, Thomas. 1888–1906. *Summa Theologiae*. Rome: Leonina.

Lord Herbert of Cherbury, Edward. 1645/1966. *De veritate*. Reprint ed. Gawlick, Stuttgart – Bad Cannstatt: Friedrich Fromann.

Grotius, Hugo. 1640. *De Iure Belli ac Pacis*. Paris.

Hobbes, Thomas. 1651. *De Cive*. London

Pufendorf, Samuel. 1997. *De officio hominis et civis*. Gesammelte Werke, Band 2. Edited by Gerald Hartung. Berlin: Akademie Verlag.

Pufendorf, Samuel. 1998. *De iure naturae et gentium*. Gesammelte Werke, Band 4. Edited by Frank Böhlting. Berlin: Akademie Verlag.

Suarez: Francesco. 1872. *Tractatus de legibus et legislatore*. Naples: Fibreniani.

### Secondary Sources

Kerting, Wolfgang. 1989. “Pflichtenlehre.” In *Historisches Wörterbuch der Philosophie*, edited by Joachim Ritter, vol. 7., 456–8. Darmstadt: Wissenschaftliche Buchgesellschaft.

# Francisco Suárez and Modern Rationalist Natural Law Theories

## Introduction<sup>1</sup>

Francisco Suárez' theory of natural law has always been the object of radically divergent interpretations. As a general rule, those commentators who blame or praise Suárez for departing from Aquinas and adopting nominalist philosophical principles usually regard him as a voluntarist (see, for example, Villey 1975, 368–95, Finnis 1980, II.6, XI.8–9, Farrell 1930, Westerman 1998, chapter 3, Bastit 1990, pt. 3, André-Vincent 1963, Moreau 1979, Delos 1950, chapter 6), while scholars seeing in him a creative innovator inside the Thomist camp are inclined to consider him as a kind of rationalist who clearly rejected the voluntarism of Ockham and Scotus (see, for example, Crowe 1976, Copleston 2003, chapter 23, Skinner 1978, chapter 5, Gierke 1900, chapter 9, Rommen 1926, Jarlot 1949, Blic 1930, Jombart 1932). A third type of interpretation suggests that Suárez deviated from the Thomistic natural law tradition in the opposite direction, so that he prepared the way for modern rationalism and secularised natural law theory (see Welzel 1980, 97–9, Wilenius 1963, 56–63, Courtine 1999, Gordley 2012).<sup>2</sup> This paper is mainly intended to refute this third interpretation of Suárez.

---

1 In referring to primary sources, I will use the following abbreviations: *AS* = Thomas Aquinas, *Scriptum super libros Sententiarum*, *BS* = Gabriel Biel, *Collectorium circa quattuor libros Sententiarum*, *CD* = Gabriel Vázquez, *Commentariorum ac disputationum in primam secundae Sancti Thomae*, *D* = William Ockham, *Dialogus*, *DL* = Francisco Suárez, *De legibus ac Deo legislatore*, *IBP* = Hugo Grotius, *De iure belli ac pacis*, *LA* = Augustine, *De libero arbitrio*, *QS* = William Ockham, *Quodlibeta septem*, *RS* = Gregory of Rimini, *Lectura super primum et secundum Sententiarum*, *SE* = Thomas Aquinas, *Sententia libri Ethicorum*, *ST* = Thomas Aquinas, *Summa theologiae*.

2 Michel Villey, John Finnis and Pauline C. Westerman accuse Suárez of voluntarist and (modern) rationalist tendencies at the same time. See Villey 1975, 384–86, Finnis 1980, 45, 338, 350, Westerman 1998, 78, 102, 108–10.

# Suárez's Rationalism

Following Aquinas (*ST I–II* q. 91 a. 2 co.), Suárez conceives of natural law as the participation of eternal law in rational beings. Eternal law and natural law differ as *lex per essentiam* and *lex per participationem*, or as “law as it exists in the lawgiver” and “law as it is in the subject” (*DL* 2, Intr.). In his discussion of the general concept of law, Suárez describes “law as in the subject” as pertaining to the intellectual nature, asserting that only rational creatures can be governed by law, whereas irrational beings, lacking reason and free will, are not capable of participating in it (*DL* 1.4.2, *DL* 1.1.2, *DL* 1.3.14). Suárez applies this principle emphatically to every kind of law. In this respect, he seems to be more consistently rationalist than Saint Thomas, who willingly incorporates Ulpian’s definition (“*quod natura omnia animalia docuit*”) into his natural law theory (see *ST I–II* q. 94 a. 2 co., *ST I–II* q. 95 a. 4 ad 1, *ST II–II* q. 57 a. 3 co., *AS IV* d. 33 q. 1 a. 1 ad 4, *SE V* l. 12 n. 4). Likewise, he is empathic in separating natural law from human will and attaching it to right reason. Deploying the Thomist language of *dominium sui*, Suárez argues that as

the exercise of dominion and the function of ruling are characteristic of law, and in man these functions are to be attributed to right reason, [...] the natural law must be constituted in the reason, as in the immediate and intrinsic rule of human actions.<sup>3</sup>

Suárez further strengthens the rationalist character of Thomistic natural law theory by extending natural law to the conclusions deduced from the primary, *per se nota* principles (Welzel 1980, 98). What is more, he elevates them to the same level of validity (and immutability) as that of the first principles.<sup>4</sup> Suárez surpasses the rationalism of Aquinas in another aspect, too. As a consequence of the certainty and necessary truth of the conclusions derived from general principles, he does not allow the slightest change in the precepts of natural law.

The natural law cannot of itself lapse or suffer change, whether in its entirety, or in its individual precepts – he declares categorically – since it is an intrinsic property which flows of necessity from human nature as such.<sup>5</sup>

- 
- 3 *DL* 2.5.12: “proprium est legis dominari et regere. Sed hoc tribuendum est rectae rationi in homine [...] in ratione est lex naturalis constituenda tanquam in proxima regula intrinseca humanarum actionum.”
  - 4 In doing so, Suárez manifestly contradicts the opinion of Aquinas that only the general principles of natural law are necessarily true, while the conclusions derived therefrom are variable and uncertain. Cf. *ST I–II* q. 94 aa. 4–5.
  - 5 *DL* 2.13.2: “legem naturalem per seipsam desinere non posse vel mutari, neque in universali neque in particulari [...]. Prout est in homine mutari non potest, quia est intrinseca proprietas necessario fluens ex tali natura, qua talis est.”

This does not mean that, unlike Aquinas, he does not take at all into consideration the contingency of human affairs (Westerman 1998, 109); but he conceives of the precepts of natural law as containing in themselves (at least implicitly) the conditions in which they should be applied (see *DL* 2.13.7, *DL* 2.13.9). That leaves, of course, a much lesser role for prudence than it has in the moral theory of Aquinas (see Treloar 1991). So Suárez appears to commit the “hubris” – characteristic of many later seventeenth- and eighteenth-century natural law theories – of laying down an all-encompassing, inflexible code of natural precepts. With this thesis Suárez incurs the criticisms of Villey, Welzel and Gordley, describing him as a rigid formalist who fossilized natural law (Villey 1975, 387–9, Welzel 1980, 98–9, Gordley 2012, 221–3). Nevertheless, this interpretation is only partly true. Suárez distinguishes two fundamentally different kinds of natural law: “perceptive” and “permissive” (*DL* 2.14.6, *DL* 2.14.14, *DL* 2.14.19, *DL* 2.18.2). And formalism is present only in the former type; “permissive” natural law, on the other hand, defines an area of human freedom and autonomy, where Suárez does justice to the variability of human conditions (Westerman 1998, 114, Wilenius 1963, 63, Tierney 2014, 204–9).

Suárez raises the crucial question of the *ratio formalis* (formal basis) of natural law. Is it rational nature itself to which human actions may be found to be appropriate or, on the contrary, inappropriate? Or is it rather rational nature understood as the faculty of judging such conformity or lack of conformity? For him, only the second answer is acceptable. As both views might be justified on Thomistic grounds (see Farrell 1930, 82–91), this question might perhaps seem purely terminological and, to be sure, somewhat artificial at first sight, but, as Michael Bertram Crowe rightly stresses, it is much more than a *lis de verbis*, inasmuch as Suárez treats this question as constituting a part of a more general and fundamental controversy, dividing intellectualists and voluntarists, as to whether natural law should be understood as a *lex indicativa* or a *lex praeceptiva* (Crowe 1977, 216–7). Suárez provides a correct summary of this complex debate, which I will quote here – for the sake of intelligibility – at full length.

On this point, the first opinion which we shall discuss is that the natural law is not a prescriptive law, properly so-called, since it is not the expression of the will of some superior; but that, on the contrary, it is a law indicating what should be done, and what should be avoided, what of its own nature is intrinsically good and necessary, and what is intrinsically evil. Thus many writers distinguish between two kinds of law, the one indicative, the other prescriptive, and hold that the natural law is law in the first sense, not in the second. This is the view expressed by Gregory of Rimini [...], who refers to Hugh of St. Victor [...], and who is followed by Gabriel Biel [...], Jacques Almain [...] and Antonio de Córdoba [...].

Accordingly, it seems that these authors would grant that the natural law is not derived from God as a Lawgiver, since it does not depend upon His will, and since, in consequence, God does not, by virtue of that law, act as a superior who lays down commands or prohibitions. Indeed, on the contrary, Gregory, whom the others follow, says that even if God did not exist, or if He did not make use of

reason, or if He did not judge of things correctly, nevertheless, if the same dictates of right reason dwelt within man, constantly assuring him, for example, that lying is evil, those dictates would still have the same legal character which they actually possess, because they would constitute a law pointing out the evil that exists intrinsically in the object.

The second opinion, diametrically opposed to the first, is that the natural law consists entirely in a divine command or prohibition proceeding from the will of God as the Author and Ruler of Nature; that, consequently, this law as it exists in God is none other than the eternal law in its capacity of commanding or prohibiting with respect to a given matter; and that, on the other hand, this same natural law, as it dwells within ourselves, is the judgment of reason, in that it reveals to us God's will as to what must be done or avoided in relation to those things which are consonant with natural reason.

This is the view one ascribes to William Ockham [...], inasmuch as he says that no act is evil save in so far as it is forbidden by God, and which could not become good if commanded by God, and conversely; whence he assumes that the whole natural law consists of divine precepts laid down by God, and susceptible of abrogation or alteration by Him. And if someone insists that such a law would be not natural but positive, the reply is that it is called natural because of its congruity with the nature of things and not with the implication that it was not externally enacted by the command of God. Jean Gerson also inclines to this opinion [...] Pierre d'Ailly [...], too, defends this view at length [...]. The same opinion is supported at length by Andreas de Novocastro.

These authorities also add that the whole basis of good and evil in matters pertaining to the law of nature is in God's will, and not in a judgment of reason, even on the part of God Himself, nor in the very things which are prescribed or forbidden by that law.<sup>6</sup>

---

6 *DL* 2.6.3–4: “In hac re prima sententia est legem naturalem non esse legem praecipientem proprie, quia non est signum voluntatis alicuius superioris, sed esse legem indicantem quid agendum vel cavendum sit, quid natura sua intrinsece bonum ac necessarium vel intrinsece malum sit. Atque ita multi distinguunt duplicem legem: unam indicantem, aliam praecipientem, et legem naturalem dicunt esse legem priori modo, non posteriori. Ita Gregorius [...], qui refert Hugonem de Sancto Victore [...] Sequitur Gabriel [...], Almainus [...], Corduba [...]

“Atque hi auctores consequenter videntur esse concessuri legem naturalem non esse a Deo ut a legislatore, quia non pendet ex voluntate Dei, et ita ex vi illius non se gerit Deus ut superior praecipiens aut prohibens. Immo ait Gregorius, quem caeteri secuti sunt, licet Deus non esset vel non uteretur ratione vel non recte de rebus iudicaret, si in homine esset idem dictamen rectae rationis dictantis v. g. malum esse mentiri, illud habiturum eandem rationem legis quam nunc habet, quia esset lex ostensiva malitiae, quae in obiecto ab intrinseco existit.

“Secunda sententia, huic extreme contraria, est legem naturalem omnino positam esse in divino imperio vel prohibitionem procedente a voluntate Dei ut auctore et gubernatore naturae, et consequenter hanc legem, ut est in Deo, nihil aliud esse quam legem aeternam ut praecipientem vel prohibentem

The distinction between indicative law and prescriptive law can be traced back to the fourteenth century. It was introduced by the Ockhamist-Augustinian theologian Gregory of Rimini in his *Commentary on the Sentences of Peter Lombard* (RS II dd. 34–7 q. 1 a. 2). The other authors cited in the text are also fourteenth- to sixteenth-century thinkers. This is a clear sign of the fact that not only the distinction itself, but also the whole intellectualist-voluntarist debate dates back only to the fourteenth century, and not before. The controversy was essentially about divine rationality and freedom, and it was the appearance of the new, voluntarist concept of law at the turn of the thirteenth and fourteenth centuries that sparked it off. While according to Aquinas's conception of eternal law there may be order in God's mind without restraining divine freedom, the concept of law as command seemed to necessitate a firm choice between a determinist and an indeterminist view of God: God is either merely a teacher of the law of nature, Himself subject to and bound by that law, or, just the opposite, a legislator acting as an arbitrary, omnipotent sovereign (Westerman 1998, 85–6, 92). Suárez appears to think, quite rightly, that the question is not adequately posed in this manner, so he does not accept either of the two opinions, and seeks instead a *via media* founded on Saint Thomas's natural law theory.

## The Critique of Intellectualism and Voluntarism

Suárez first sets out to refute the intellectualist position. It is important to note that practically all the theologians enumerated by Suárez in this connection were (at least partly) Ockhamists and hence (more or less) voluntarists, who adopted certain essentialist viewpoints in order to differentiate themselves from the more robust voluntarism of Ockham, Gerson and d'Ailly. Gregory of Rimini and Gabriel Biel in his wake differentiated between *lex indicativa* and *lex praeceptiva* just with the purpose of counterbalancing or reconciling the voluntarist stand that it is God's will that determines what is good and evil with the

---

in tali materia. In nobis vero hanc legem naturalem esse iudicium rationis, quatenus nobis significat voluntatem Dei de agendis et vitandis circa ea quae rationi naturali consentanea sunt.

“Ita sumitur ex Ochamo [...] quatenus dicit nullum esse actum malum, nisi quatenus a Deo prohibitus est, et qui non possit fieri bonus si a Deo praecipiat, et e converso. Unde supponit totam legem naturalem consistere in praeceptis divinis a Deo posit, quae ipse posset auferre et mutare. Quod si instet aliquis talem legem non naturalem esse sed positivam, responderet dici naturalem quia est proportionata naturis rerum, non quia non sit extrinsecus a Deo posita. Et in hanc sententiam inclinat Gerson [...]. Et hanc sententiam defendunt late Petrus Alliacus [...]. Idem latissime Andreas de Novo Castro [...].

“Qui etiam addunt totam rationem boni et mali in rebus ad legem naturae pertinentibus positam esse in voluntate Dei et non in iudicio rationis, etiam ipsius Dei, neque in rebus ipsis quae per talem legem vetantur aut praecipuntur.”

rationalist view of good and evil as grounded in the nature of things (Crowe 1976, 398). That is why Gregory stressed so much that sin is sin because it is against divine reason insofar as it is right, rather than insofar as it is divine; moreover, he added,

if, under the impossible hypothesis that [*per impossibile*] divine reason or God Himself did not exist, or His reason should err, still if someone were to act against angelic or human right reason, or any other possible kind of right reason, he would sin.<sup>7</sup>

Thus, paradoxically enough, it was a “rationalist” or moderate voluntarist current inside the nominalist camp that led gradually to intellectualism and in the end to the “*etiamsi daremus*” hypothesis of Hugo Grotius (St. Leger 1962, 124, cf. *IBP* Prol. 11).

Seeing that the above-mentioned authors also aim, in their own way, at a synthesis of voluntarism and rationalism, their theoretical position seems by no means so far removed from that of Suárez as it might appear from the *De Legibus*. Does this mean that Suárez is merely tilting at windmills? Not at all. It is generally true that the theologians of the Counter-Reformation, in contrast with the nominalist and voluntarist tendencies inherent in Protestant thought, were inclined towards intellectualism. Indeed, the reassertion of the predominance of reason in moral and legal philosophy was an integrant part of the sixteenth–seventeenth-century scholastic revival (St. Leger 1962, 93). In this way, the Second Scholasticism played an important role in the revitalisation of classical Aristotelian-Thomist natural law doctrine. However, certain Thomists, in their zeal to oppose Protestant voluntarism, leaned towards extreme objectivism and rationalism in law and morals. This tendency culminated in the *oeuvre* of Suárez’s fellow Jesuit Gabriel Vázquez.<sup>8</sup> Hans Welzel points out that by incorporating the phrase “*vel non recte de rebus iudicaret*” (or if He did not judge of things correctly) into the description of the intellectualist conception of natural law, Suárez made an unequivocal allusion to his Jesuit rival (Welzel 1980, 97), who had affirmed in his Commentary on the *Prima Secundae* of the *Summa theologiae* that “if we should concede, which is indeed impossible, that *God did not judge as He does now*, and if there remained in us the use of reason, sin would also remain.”<sup>9</sup> The ground for this statement is that sin is evil of itself, prior to any external prohibition, even to the judgment or will of God (*CD* d. 97 c. 1 n. 2). From this allusion it seems quite obvious that Suárez’s criticism is directed as much, if not more, against Vázquez and other contemporary exponents of extreme intellectualism than against Gregory of Rimini and Biel. It is worth

7 *RS* II dd. 34–7 q. 1 a. 2: “si per impossibile ratio divina sive deus ipse non esset aut ratio illa esset errans, adhuc, si quis ageret contra rectam rationem angelicam vel humanam aut aliam aliquam, si qua esset, peccaret.” Biel reiterated almost verbatim Gregory’s statement. See *BS* II d. 35 q. 1 a. 1.

8 As a general rule, Jesuits were more inclined to extreme essentialism than Dominicans. Vitoria and Soto, for instance, were much more moderate in this respect.

9 *CD* d. 97 c. 1 n. 3: “si concessio impossibili intelligeremus *Deum non ita iudicare*, et manere in nobis usum rationis, maneret etiam peccatum” (emphasis added).



mentioning here that in an earlier phase of his scientific career, Suárez too tended towards extreme essentialism (Suñer 1974, xlii–vi).<sup>10</sup>

Here we have to return for a while to the problem of the *ratio formalis* of natural law. In Suárez's age, the major proponent of the view that natural law should be identified with rational nature as such was none other than Gabriel Vázquez.<sup>11</sup> Vázquez located the formal basis of natural law in human rational nature itself, rather than in the judgment of reason in order to eliminate all subjective elements from the concept of *ratio* (Welzel 1980, 95–6). This opinion is unacceptable for Suárez. Not that he questions the doctrine of *perseitas boni* which assumes the intrinsic goodness (or badness) of actions. Just the contrary! He willingly accepts the idea that rational nature is the foundation of the objective goodness of moral actions; but on the other hand, he dismisses the view that for that reason it can be termed law. Suárez admits that human rational nature is rightly considered as a measure or standard for law but not that '*lex*' can be predicated of it. For "rational nature itself, strictly viewed in its essential aspect, neither gives commands, nor makes evident the rectitude or turpitude of anything".<sup>12</sup>

In Suárez's opinion, by equating *lex naturalis* with rational nature, Vázquez defends a non-legal conception of natural law, and this holds in general for all (extreme) intellectualists. Moreover, this conception would lead to absurd conclusions. If rational nature or the judgment of right reason alone sufficed to constitute law, then God would have His own natural law, binding and obligatory on Him.<sup>13</sup>

God Himself would be subject to a natural law relating to His will; since even in God, an intellectual act of judgment logically precedes an act of His will, a judgment indicating that lying is wicked, that to keep one's promises is wholly right and necessary, and so forth; and therefore, if such an act of the intellect is sufficient to constitute the essence of law, then there will be a true natural law, even with respect to God Himself. For in such a case, the fact that God has no superior, will not serve as an objection, since the natural law is not imposed by any superior.<sup>14</sup>

---

10 Hugo Grotius, as is well known, made an intellectual move in the opposite direction: he first advocated a voluntarist doctrine of natural law in *De iure praedae*, and then opted for thoroughgoing rationalism in *De iure belli ac pacis*.

11 See, for example, *CD* d. 150 c. 3 n. 23: "Prima igitur lex naturalis in creatura rationali est ipsamet natura, quatenus, rationalis, quia haec est prima regula boni et mali."

12 *DL* 2.5.6: "non omne id quod est fundamentum honestatis seu rectitudinis actus lege praecepti vel quod est fundamentum turpitudinis actus lege prohibiti, potest dici lex. Ergo licet natura rationalis sit fundamentum honestatis obiectivae actuum moralium humanorum, non ideo dici potest lex. Et eadem ratione, quamvis dicatur mensura, non ideo recte concluditur quod sit lex, quia mensura latius patet quam lex." *DL* 2.5.5: "natura ipsa rationalis praecise spectata, ut talis essentia est, nec praecipit, nec ostendit honestatem aut malitiam, nec dirigit aut illuminat, nec alium proprium effectum legis habet."

13 *DL* 2.5.7: "Praeterea possumus ab inconvenientibus argumentari. Unum est, quia sequitur non minus proprie habere Deum suam legem naturalem quae ipsum liget et obliget, quam homines."

14 *DL* 2.6.6: "etiam Deus haberet legem sibi naturalem respectu suae voluntatis, quia etiam in Deo ad voluntatem antecedit secundum rationem iudicium mentis, indicans mentiri esse malum, servare



Nevertheless, this is not the single or greatest error that intellectualism commits in connection with natural law according to Suárez. Beside the fact that the intellectualists undermine the legal character of natural law they also make it doubtful that it is truly divine law. For the intellectualist standpoint entails that

the precepts of the natural law are not from God, inasmuch as they are characterized by a necessary goodness, and inasmuch as that condition of necessary goodness, which is in rational nature (by reason of which that nature is the measure of such goodness), does not depend upon God for its rational basis, although its actual existence does depend upon Him. [...] Hence, natural law is prior to the divine judgment and the divine will of God; and therefore, natural law does not have God for its author, but necessarily dwells within rational nature in that matter, in such fashion that it is inherently endowed with this essence, and no other.<sup>15</sup>

This passage, I think, clearly shows that Suárez is fully aware of the possible secularist implications of a full-blown rationalist conception of natural law.<sup>16</sup> That is why he got so scared of Vázquez's natural law doctrine (Welzel 1980, 97). This danger was already inherent in the "*etiamsi daremus*" hypothesis of Gregory of Rimini and Gabriel Biel – even though if this was very far from their original intentions. As James St. Leger rightly stresses, they considered the supposition of the non-existence of God simply as an impossible condition, a condition contrary to fact, and "the only purpose of this hypothesis was to bring into bold relief the rational character of natural law as opposed to the voluntarism of authors who linked the natural law exclusively to a command of the divine will." (St. Leger 1962, 123). *Mutatis mutandis*, this is also true of Grotius, who took over this medieval commonplace with the intention of underlining the rationality and immutability of the moral order, and not with the purpose of separating law from theology or of constructing a secularised theory of natural law (Crowe 1976, 381, 405, Chroust 1943, 126).

---

promissum esse omnino rectum et necessarium. Si ergo hoc satis est ad rationem legis, etiam in Deo erit vera lex naturalis. Quia tunc non obstat quod Deus non habeat superiorem, quia lex naturalis non imponitur ab aliquo superiore." Likewise, a rational judgment of an equal, of an inferior or of a teacher showing the nature of a given action, adds Suárez, would be "law" in the proper sense. Such a conclusion would be manifestly absurd, too.

- 15 *DL* 2.5.8: "Deinde sequitur legem naturalem non esse legem divinam, neque esse ex Deo. Probatur sequela, quia iuxta illam sententiam praecepta huius legis non sunt ex Deo quatenus necessariam honestatem habent, et illa conditio quae est in natura rationali, ratione cuius est mensura illius honestatis, non pendet a Deo in ratione, licet pendeat in existentia. [...] Ergo lex naturalis praecedit iudicium et voluntatem Dei. Ergo non habet auctorem Deum, sed per se inest tali naturae eo modo quo de se habet ut sit talis essentiae et non alterius."
- 16 Reijo Wilenius argues that Suárez makes natural law and the moral order "autonomous, independent of God's will". It seems to me that Wilenius is committing here the very mistake he warns against a bit later: "One is easily misled in Suárez's works by the fact that he puts forth with the utmost care, and as if they were his own, opinions which he later refutes." See Wilenius (1963, 59–60, 60 n. 3).

Suárez is no less critical of the voluntarist conception of natural law. He is convinced that if intellectualism denies the prescriptive and hence the legal character of natural law, then voluntarism precludes its “naturalness”, since it bases natural law on arbitrary divine fiat. For Ockham, as Suárez not altogether correctly reads him, divine volition is the sole source of good and evil.<sup>17</sup> The *doctor eximius* dismisses this view as “false and absurd” (DL 2.15.4). He tackles the questions of hatred of God and adultery that after Ockham became the nerve-points of the intellectualist-voluntarist controversy, and rejects forcefully the answers given to them by the nominalist theologian. He lays down as “an axiom common to the theologians that certain evils are prohibited because they are evil.”<sup>18</sup> He traces this axiom back to Saint Augustine, saying (through Evodius) in *De libero arbitrio* (LA 1.3) that adultery is not an evil because prohibited by law (*malum quia prohibitum*), but it is so prohibited because it is evil (*malum per se*).<sup>19</sup> Furthermore, he recalls the metaphysical principle that the nature of things, their essence, is immutable. Some human acts are intrinsically, by their very nature good or bad. If this were not the case, then it would be possible even for hatred of God to become righteous and allowed by Him, which would be plain nonsense (DL 2.6.11, DL 2.15.4).

## The Suárezian via media

I have noted earlier that Suárez searches for a Thomist middle course that avoids both the Scylla of (extreme) intellectualism and the Charybdis of voluntarism. We have seen above his objections to these two extremes. But what does his own solution consist in? Suárez suggests that the natural law is a *lex indicativa* and a *lex praeceptiva* at the same time:

natural law, as it exists in man, does not merely indicate what is evil, but actually obliges us to avoid the same; [...] it consequently does not merely point out the natural disharmony of a particular act or object with rational nature, but is also a manifestation of the divine will prohibiting that act or object.<sup>20</sup>

---

17 Suárez disregards the rationalist side of Ockham’s moral philosophy. He ignores, for instance, Ockham’s claim that moral doctrine is twofold, and demonstrative, non-positive moral science directs human acts apart from any command or precept of a superior. See QS II q. 14.

18 DL 2.6.11: “Et quoad priorem partem colligitur ex illo communi axioma theologorum: quaedam mala esse prohibita quia mala.”

19 In order to prove his thesis, Suárez (DL 2.6.18) invokes the authority of Aquinas as well, according to whom (ST I–II q. 71. a. 6 ad 4) a sin is contrary to the natural law “precisely because it is inordinate”.

20 DL 2.6.13: “fit legem naturalem, prout in nobis est, non tantum esse indicantem malum, sed etiam obligantem ad cavendum illud, [...] subinde non solum repraesentare naturalem disconvenientiam talis actus vel obiecti cum rationali natura, sed etiam esse signum divinae voluntatis vetantis illud.”

On the one hand, the divine precept or prohibition does not constitute the whole reason of the good or evil involved in the observance or transgression of natural law. On the contrary, God's will necessarily presupposes in the object of the act concerned the existence of an intrinsic harmony or disharmony with rational nature and with its proper end (*perseitas boni et mali*).<sup>21</sup> Thus the natural law is indeed "natural". On the other hand, natural reason indicates not only that something is in itself good or evil, but also that it is in conformity with the divine will that the good should be done and the evil avoided.<sup>22</sup> The divine volition attaches to the goodness or badness inherent in the relevant acts an obligation derived from divine law:

all things which are declared evil by the natural law are forbidden by God, by a special command and by that will which binds and obliges us, through the force of His authority, to obey those natural precepts; therefore, the natural law is truly prescriptive law, that is to say, one which contains true precepts [...] the natural law is truly and properly divine law, of which God is the Author.<sup>23</sup>

Suárez here has to face a problem analogous to but not identical with that of the "*etiamsi daremus*" hypothesis: if God were not to issue the prohibitions and commands of natural law, would lying nevertheless be evil and a sin, and respecting one's parents good? Suárez gives a complex answer to the question. First, disaccord with right reason is in itself, apart from its relation to law, a moral evil and a sin.<sup>24</sup> Secondly, a sin forbidden by God

is also characterized by a special depravity which it would not possess if the divine prohibition had not intervened, and it is in view of this depravity that the character of sin considered theologically becomes complete.<sup>25</sup>

---

21 DL 2.6.11: "Haec Dei voluntas, prohibitio aut praeceptio non est tota ratio bonitatis et malitiae quae est in observatione vel transgressione legis naturalis, sed supponit in ipsis actibus necessariam quamdam honestatem vel turpitudinem."

22 DL 2.6.8: "ratio naturalis quae indicat quid sit per se malum vel bonum homini, consequenter indicat esse secundum divinam voluntatem ut unum fiat et aliud vitetur."

23 Ibid.: "omnia quae lex naturalis dictat esse mala, prohibentur a Deo speciali praecepto et voluntate, quae vult nos teneri et obligari vi auctoritatis eius ad illa servanda. Ergo lex naturalis est proprie lex praeceptiva seu insinuativa proprii praecepti." DL 2.6.13: "legem naturalem esse veram ac propriam legem divinam, cuius legislator est Deus."

24 DL 2.6.17: "Respondeo igitur in actu humano esse aliquam bonitatem vel malitiam ex vi obiecti prae-cise spectati, ut est consonum vel dissonum rationi rectae et secundum eam posse denominari et malum et peccatum et culpabilem secundum illos respectus, seclusa habitudine ad propriam legem."

25 DL 2.6.18: "Quae ratio potius videtur probare esse prohibitum quia malum, quam e converso. Quod verum est loquendo de malitia moralis inordinationis, tamen ratione illius addita est lex aeterna et divina prohibitio ad quam habet tale peccatum specialem repugnantiam; et consequenter inde habet specialem deordinationem quam non haberet si prohibitio divina non intervenisset, per quam deordinationem completur ratio peccati theologicè sumpti."

Thirdly, from Aquinas's dictum that God "would deny Himself if He were to do away with the very order of His own justice",<sup>26</sup> Suárez deduces the proposition that God cannot but prohibit what is evil *per se*:

whatever is contrary to right reason is displeasing to God, and the opposite is pleasing Him; for the will of God is supremely just, and therefore, that which is evil cannot fail to displease Him, nor can that which is righteous fail to please Him, inasmuch as God's will cannot be irrational.<sup>27</sup>

Consequently, God cannot grant any dispensation from the precepts of natural law containing all the ten commandments of the Decalogue (*DL* 2.15.3-12, *DL* 2.15.16, *DL* 2.15.26).<sup>28</sup>

Suárez takes great care to embed his voluntarist concept of law into an objectivist, rationalist framework based on a metaphysical view of human nature (Abril 1974, lxxxix). He underlines that "natural law, in all its precepts, relates to the natural qualities of man",<sup>29</sup> and follows Aquinas in linking natural law to the order of natural inclinations and the teleology of human nature:

Saint Thomas [...] traces this variety in the natural precepts to the varied natural inclinations of man. For man is, as it were, an individual entity and as such has an inclination to preserve his own being, and to safeguard his own welfare; he is also a being corruptible – that is to say mortal – and as such is inclined towards the preservation of the species, and towards the actions necessary to that end; and finally, he is a rational being and as such is suited for immortality and spiritual perfection, as well as for communication with God and social intercourse with rational creatures. Hence, the natural law brings man to perfection with regard to every one of his tendencies [...] all these precepts proceed, by a certain necessity, from nature, and from God as the Author of nature, and all tend to the same end, which is undoubtedly the due preservation and natural perfection or felicity of human nature.<sup>30</sup>

---

26 *ST* I-II q. 100 a. 8 ad 2: "Negaret autem seipsum, si ipsum ordinem suae iustitiae auferret." Aquinas asserts that in connection with the question whether the precepts of the Decalogue are dispensable.

27 *DL* 2.6.5: "Et ait Deum non posse negare seipsum, et ideo non posse ordinem suae iustitiae auferre, sentiens non posse non prohibere ea quae mala sunt et contra rationem naturalem." *DL* 2.6.8: "quidquid contra rationem rectam fit, displicet Deo, et contrarium illi placet; quia cum voluntas Dei sit summe iusta, non potest illi non displicere quod turpe est, nec non placere honestum, quia voluntas Dei non potest esse irrationabilis."

28 Suárez emphatically discards the opinion of Ockham, d'Ailly and Gerson, according to which God can dispense from virtually all the precepts of the Decalogue, just as the view of Scotus and Biel that only the first table of the Decalogue does not admit of dispensation.

29 *DL* 2.14.8: "naturale ius quoad omnia praecepta sua pertinet ad naturales hominis proprietates."

30 *DL* 2.8.4: "reducit divus Thomas [...] varietatem hanc praeceptorum naturalium ad varias hominis inclinationes naturales. Est enim homo individuum quoddam ens et ut sic inclinatur ad conservandum

# Towards Secularised Rationalism?

The above theses seem to be in perfect harmony with the spirit of Aquinas and in direct opposition to the Ockhamist doctrine that God can command (or abstain from commanding) virtually anything. Accordingly, it would be a gross mistake to label Suárez without qualification a voluntarist, as Villey, Farrell and certain other scholars do. At first sight, the interpretation of Hans Welzel seems much better founded. He argues that, after all, the Suárezian middle course is nothing but a compromise, for “fundamentally, in Suárez too [just as in Vázquez], the divine will remains bound to the rational nature of things. God must forbid what is intrinsically evil and against natural reason. To the self-existing good or evil God’s will only appends the special obligation of divine law.” (Welzel 1980, 97–8). Similarly, Thomas Pink deems that Suárez “comes close to conceding the substance” of Vázquez’s intellectualist conception of natural law as “a law without a law-maker” (Pink 2012, 134–5). John Finnis, albeit emphasizing the voluntarism of Suárez’s notion of obligation, also affirms that the Spanish theologian “is brought to the brink of saying that even without reference to any divine precept, acts (or their avoidance) can be obligatory” (Finnis 1980, 350). Jean-François Courtine goes yet further, claiming that “in spite of the tirelessly reiterated criticisms against Vázquez, it is legitimate to ask whether Suárez does not concede the essence. Certainly, he does not maintain without a corrective the radical thesis that the *dictamen naturale rectae rationis* as such has the force of law; this would be to consider that man, completely rational, is a law unto himself. However, the correction made here by Suárez, i.e. the necessity of the supplement that consists in the imperative as a sign of the will, does not modify in substance the underlying thesis of autonomy.”<sup>31</sup> (Courtine 1999, 113). Courtine’s final conclusion is that in the wake of Vázquez, Suárez made an important contribution to the rationalisation and secularisation of natural law (Courtine 1999, 114).

Would the decree of God’s will really be solely a supplement to the judgment of right reason in Suárez’s natural law theory? This would imply a determinist view of God that

---

suum esse ad suam commoditatem. Est etiam ens corruptibile seu mortale et ut sic inclinatur ad conservationem speciei et ad actiones propter illam necessarias. Tandem rationalis est et ut sic capax immortalitatis et spiritualium perfectionum et communicationis cum Deo ac societatis cum rationalibus creaturis. Lex ergo naturalis perficit hominem secundum omnem inclinationem suam.” *DL* 2.7.7: “haec omnia praecepta necessitate quadam prodeunt a natura et a Deo quatenus auctor est naturae, et tendunt ad eundem finem, nimirum ad debitam conservationem et naturalem perfectionem seu felicitatem humanae naturae.”

- 31 Courtine supports his claim with the following quotation, taken from *DL* 2.5.10: “such dictates have the force of law over man, even though they may not be externally clothed in the form of written law. Therefore, these dictates constitute natural law; and accordingly, the man who is guided by them is said to be a law unto himself, since he bears law written within himself through the medium of the dictates of natural reason” [consequenter ostendit dictamen illud habere vim legis in homine, etiamsi scriptam exterius legem non habeat. Hoc ergo dictamen est lex naturalis et ratione illius dicitur homo qui illo ducitur esse sibi lex, quia in se habet scriptam legem medio dictamine naturalis rationis].

would entirely destroy God's freedom and consequently Suárez's conception of *lex aeterna* based on it. And it would contradict his deep conviction that law is an act of free will and a command of a superior<sup>32</sup> (not to mention the title of his book: *On Laws and God the Law-giver*). So either Suárez contradicts himself or Courtine's suggestion is wrong. Everything seems to turn on what Suárez exactly means by saying that "God cannot fail to prohibit that which is intrinsically evil and inordinate in rational nature".<sup>33</sup> And this in turn depends on the stance he takes in the old dispute over the absolute and ordained power of God. To be sure, Suárez disapproves Ockham's voluntarist view that God can, by virtue of His absolute power, abstain from laying down such a prohibition,<sup>34</sup> and he affirms that the divine will necessarily presupposes a dictate of the divine reason declaring that a given act is righteous or evil.<sup>35</sup> But on the other hand he repeatedly and vigorously denies that either human rational nature or the judgment of divine reason constitutes a law binding God's will. God is entirely free from law, thus what He wills is always just and fitting.<sup>36</sup>

Hence, notwithstanding any law whatsoever made by Himself for the government of Creation, God may disregard that law, making use of His absolute power, as in the distribution of rewards or punishments, and so forth; because He is not bound to the observance of law. For He is sovereign Lord and not confined within any order.<sup>37</sup>

So it is evident that by saying that God has to prohibit what is evil *per se*, Suárez by no means suggests that God is legally obliged to do so or that he issues that prohibition mechanically, without deliberation. Yet He cannot but prohibit evil, but for rather different reasons (which have to be quoted at full length here):

---

32 See, for example, *DL* 1.5.24: "law [...] is the act of a just and upright will, the act whereby a superior wills to bind an inferior to the performance of a particular deed" [legem (...) esse actus voluntatis iustae et rectae, quo superior vult inferiorem obligare ad hoc vel illud faciendum].

33 *DL* 2.6.21: "non potest Deus non prohibere id quod est intrinsece malum et inordinatum in natura rationali."

34 *DL* 2.6.20: "In quo duo possunt cogitari modi dicendi. Primus est Deum quidem posse de potentia absoluta non facere talem prohibitionem, quia non apparet implicatio contradictionis, ut videntur probare omnia quae Ockham, Gerson et alii pro sua sententia congerunt; nihilominus tamen id fieri non posse secundum legem ordinariam divinae providentiae rerum naturis consentaneam. Nam hoc ad minus probant rationes in contrarium factae pro nostra sententia et multum favent testimonia Scripturae et Patrum."

35 *DL* 2.6.13: "Unde probandum non est quod doctores posteriori loco allegati dicunt voluntatem divinam, qua lex naturalis sancitur, non supponere dictamen divinae rationis dictantis hoc esse honestum vel turpe."

36 *DL* 2.2.5: "Anselmus dicens Deum esse omnino liberum a lege et ideo quod vult, iustum et conveniens esse."

37 *DL* 2.2.6: "Unde non obstante quacumque lege a se posita circa rerum gubernationem, potest illam non servare, sua potentia absoluta utendo, ut circa praemia vel poenam retribuendam et similia, quia non obligatur ad servandam legem, quia est supremus Dominus et extra omnem ordinem."

although the divine will is absolutely free in its external actions, nevertheless, if it be assumed that this will elicits one free act, then it may be necessarily bound, in consequence, to the performance of another action. For example, if through the divine will an unconditional promise is made, that will is obliged to fulfil the promise [...]. In like manner, if it is the divine will to create the world, and to preserve the same in such a way as to fulfil a certain end, then there cannot fail to exist a providential care over that world [...]. Accordingly, assuming the existence of the will to create rational nature with sufficient knowledge for the doing of good and evil, and with sufficient divine co-operation for the performance of both, God could not have refrained from willing to forbid that a creature so endowed should commit acts intrinsically evil, nor could He have willed not to prescribe the necessary righteous acts. For just as God cannot lie, neither can He govern unwisely or unjustly; [...] absolutely speaking, God could have refrained from laying down any command or prohibition; yet, assuming that He has willed to have subjects endowed with the use of reason, He could not have failed to be their lawgiver – in those matters, at least, which are necessary to natural moral rectitude.<sup>38</sup>

This solution is slightly reminiscent of Ockham’s conception of conditional natural law,<sup>39</sup> insofar as the core of Suárez’s argument is as follows: *supposing* that God has decided to create man as a rational, free being, then he could not have abstained from commanding/forbidding him what is according/contrary to his nature and hence shown by right reason to be in itself good/evil. But this is, I think, no more than a formal resemblance. Substantially, Suárez’s argument is much closer to the Thomist view that while theoretically it is conceivable that by his absolute power God could act independently of the created order, in effect, God’s will always coincides with the order which He has established (see Pernoud 1972, 83, cf. *ST I* q. 25 a. 5 ad 1). Following his theory of ordered causes, Aquinas states that if we consider the order of nature established by God depending on the first cause, i.e. Himself,

---

38 *DL* 2.6.23: “Dico igitur ex Caietano divinam voluntatem, licet simpliciter libera sit ad extra, tamen ex suppositione unius actus liberi posse necessitari ad alium ut, si vult promittere absolute, necessitatur ad implendum promissum. [...] Et cum eadem proportione, si vult creare mundum et illum conservare in ordine ad talem finem, non potest non habere providentiam illius [...]. Ideoque supposita voluntate creandi naturam rationalem cum sufficienti cognitione ad operandum bonum et malum et cum sufficienti concursu ex parte Dei ad utrumque, non potuisse Deum non velle prohibere tali creaturae actus intrinsece malos vel nolle praecipere honestos necessarios. Quia sicut non potest Deus mentiri, ita non potest insipienter vel iniuste gubernare [...] absolute posset Deus nihil praecipere vel prohibere. Tamen ex suppositione quod voluit habere subditos ratione utentes, non potuit non esse legislator eorum saltem in his quae ad honestatem naturalem morum necessaria sunt.”

39 Ockham differentiates between three usages of the term *ius naturale*. The third meaning – what he calls “*ius naturale ex suppositione*” – contains rational answers to contingent conditions. See *D* 3.2.3.6.



God cannot do anything against this order; for, if He did so, He would act against His foreknowledge, or His will, or His goodness. But if we consider the order of things depending on any secondary cause, thus God can do something outside such order; for He is not subject to the order of secondary causes; but, on the contrary, this order is subject to Him, as proceeding from Him, not by a natural necessity, but by the choice of His own will; for He could have created another order of things.<sup>40</sup>

Although for the things already made no other order would be fitting and good, God could do other things, and impose upon them another order, as He is bound to nobody but Himself. This means, on the other hand, that He can do nothing but what is befitting to Himself and just.<sup>41</sup>

Thus Suárez agrees with Saint Thomas (and partly with Ockham) that God could have created another moral order. The act of creation is a completely free act; the divine will can freely choose between several rational plans. And after decreeing absolutely that something is to be done or to be avoided, “God is unable to act in opposition to His own decree not on account of any prohibition which the decree carries with it, but on account of the repugnant nature of that act itself”.<sup>42</sup> God could in principle rightfully do so, but this would be against His very nature. And obviously God cannot deny Himself and cannot abolish the order of His own justice:

granted that it implies not a physical contradiction (so to speak), but solely a moral one, for God to change His decree, and further, granted that once He has made a decree, it is contrary to due order that He should act in opposition thereto, nevertheless, these facts result not from any prohibition but from the intrinsic nature and essence of God [...]. For just as it is unfitting that divinity should deceive, even so it is unfitting that divinity should be inconstant.<sup>43</sup>

---

40 *ST I* q. 105 a. 6 co.: “Si ergo ordo rerum consideretur prout dependet a prima causa, sic contra rerum ordinem Deus facere non potest, sic enim si faceret, faceret contra suam praescientiam aut voluntatem aut bonitatem. Si vero consideretur rerum ordo prout dependet a qualibet secundarum causarum, sic Deus potest facere praeter ordinem rerum. Quia ordini secundarum causarum ipse non est subiectus, sed talis ordo ei subiicitur, quasi ab eo procedens non per necessitatem naturae, sed per arbitrium voluntatis, potuisset enim et alium ordinem rerum instituire.”

41 *ST I* q. 25 a. 5 ad 3: “licet istis rebus quae nunc sunt, nullus alius cursus esset bonus et conveniens, tamen Deus posset alias res facere, et alium eis imponere ordinem.” *ST I* q. 25 a. 5 ad 2: “Deus non debet aliquid alicui nisi sibi. Unde, cum dicitur quod Deus non potest facere nisi quod debet nihil aliud significatur nisi quod Deus non potest facere nisi quod ei est conveniens et iustum.”

42 *DL* 2.2.7: “Dices: si Deus, postquam decrevit absolute aliquid non facere, id ageret, inordinate faceret, et ideo id facere non potest. Ergo liberum decretum Dei habet vim positivae legis respectu voluntatis eius, ut non possit honeste facere quod per se ac remoto illo decreto libere facere potuisset. Responde Deum non posse facere contra suum decretum, non propter prohibitionem quam decretum inducat sed propter repugnantiam ipsius rei.”

43 *Ibid.*: “esto non implicaret contradictionem physicam (ut sic dicam) mutare Deum decretum suum, sed tantum morale ac subinde posito uno decreto esse inordinatum agere contra illud. Nihilominus id non oriri ex prohibitione, sed ex intrinseca natura et essentia Dei [...]. Quia sicut non decet divinitatem fallere, ita nec inconstantem esse.”



The fact that Suárez formulates the above opinion in his discussion of eternal law can be a perfect illustration why he holds to this traditional idea, which he conceives, in contrast with its traditional scholastic meaning, not as a norm above the divine volition but as a free expression of God's will (see *DL* 2.2.9, *DL* 2.3.4, *DL* 2.3.6). Eternal law and creation are absolutely free acts of God, whereas all His subsequent acts – including the precepts of natural law – are only relatively free, being bound in consequence of them.

## Conclusion

Francisco Suárez faced the intriguing but difficult challenge of accommodating the late medieval voluntarist concept of law, together with all the problems it entails, into his predominantly Thomist and rationalist legal philosophy. As it seems to me, he succeeded in elaborating an equilibrated and viable synthesis of essentialism and voluntarism. While on the one hand he restored the “naturalness” and rationality of natural law on the basis of the teleology and *perseitas moralis* of human acts, on the other hand the voluntarist elements of the Suárezian synthesis, above all his conception of “*lex aeterna*”, did not leave space for any kind of secularised rationalism. In this way he was able to avoid the trap of both voluntarism and extreme rationalism, into which most modern natural law theorists later fell.

## Bibliography

### Primary Sources

- AS Aquinas, Thomas (Saint). 1929–1947. *Scriptum super libros Sententiarum*. Edited by Pierre Mandonnet and Marie-Fabien Moos, 4 vols. Paris: Lethielleux.
- BS Biel, Gabriel. 1973–1992. *Collectorium circa quattuor libros Sententiarum*. Edited by Wilfrid Werbeck and Udo Hofmann, 6 vols. Tübingen: Mohr.
- CD Vázquez, Gabriel. 1631. *Commentariorum ac disputationum in primam secundae Sancti Thomae*. 2 vols. Lyon: Cardon.
- D Ockham, William. *Dialogus: Latin Text and English Translation*. Edited and translated by John Kilcullen et al. (in progress). <http://www.britac.ac.uk/pubs/dialogus/ockdial.html>
- DL Suárez, Francisco. 1971–1981. *De legibus*. Edited by Luciano Pereña et al, 8 vols. Madrid: Consejo Superior de Investigaciones Científicas.
- Suárez, Francisco. 1944. “On Laws and God the Lawgiver” (selections). In *Selections from Three Works of Francisco Suárez, S. J., An English Version of the Texts*, translated by Gwladys L. Williams, Ammi Brown and John Waldron, vol. 2., 1–646. Oxford: Clarendon Press.

- IBP Grotius, Hugo. 2006 [1625]. *De iure belli ac pacis*. Paris: Buon. Reprint, Hildesheim: Olms.
- LA Augustine (Saint). 1845. “De libero arbitrio.” In *Patrologia Latina*, edited by Jacques-Paul Migne, vol. 32, 1221–1310. Paris: Garnier.
- QS Ockham, William. 1980. “Quodlibeta septem.” In *Opera theological*, edited by Joseph C. Wey, vol. 9, St. Bonaventure: St. Bonaventure University, Franciscan Institute.
- RS Rimini, Gregory of. 1979–1987. *Lectura super primum et secundum Sententiarum*. Edited by A. Damasus Trapp and Venício Marcolino, 7 vols. Berlin: De Gruyter.
- SE Aquinas, Thomas (Saint). 1969. “Sententia libri Ethicorum.” In *Opera omnia*. Leonine edition, vols. 47/1–2. Rome: Sancta Sabina.
- ST Aquinas, Thomas (Saint). 1888–1906. “Summa theologiae.” In *Opera omnia*. Leonine edition, vols. 4–12. Rome: Typographia Polyglotta.  
 Aquinas, Thomas (Saint). 1920–25. *Summa theologiae*. Translated by Fathers of the English Dominican Province, 2nd ed., 22 vols. London: Burns, Oates & Washbourne.

## Secondary Sources

- Abril, Vidal. 1974. “Perspectivas del iusnaturalismo suareciano.” In Francisco Suárez. *De legibus*, edited by Luciano Pereña et al., vol. 3.; *De legibus (II 1–12): De lege naturali*, edited by Luciano Pereña and Vidal Abril, lvi–lxxxvi. Madrid: Consejo Superior de Investigaciones Científicas.
- André-Vincent, Philippe-Ignace. 1963. “La notion moderne de droit naturel et le volontarisme (de Vitoria et Suarez à Rousseau).” *Archives de philosophie du droit* 8: 237–59.
- Bastit, Michel. 1990. *Naissance de la loi moderne: La pensée de la loi de Saint Thomas à Suarez*. Paris: Presses Universitaires de France.
- Blic, Jacques de. 1930. “Le volontarisme juridique chez Suarez?” *Revue de philosophie* 1: 213–30.
- Chroust, Anton-Hermann. 1943. “Hugo Grotius and the Scholastic Natural Law Tradition.” *The New Scholasticism* 17(2): 101–33.
- Copleston, Frederick. 2003. *A History of Philosophy. Late Medieval and Renaissance Philosophy*. Vol. 3. London: Continuum.
- Courtine, Jean-François. 1999. “La raison et l’empire de la loi.” In *Nature et empire de la loi: Études suaréziennes*, 91–114. Paris: Vrin.
- Crowe, Michael Bertram. 1977. *The Changing Profile of the Natural Law*. The Hague: Nijhoff.
- Crowe, Michael Bertram. 1976. “The ‘Impious Hypothesis’: A Paradox in Hugo Grotius?” *Tijdschrift voor Filosofie* 38(3): 379–410.
- Delos, Joseph Thomas. 1950. *La société internationale et les principes du droit public*. 2nd ed. Paris: Pedone.
- Farrell, Walter. 1930. *The Natural Moral Law according to St. Thomas and Suarez*. Ditchling: St. Dominic’s Press.
- Finnis, John. 1980. *Natural Law and Natural Rights*. Oxford: Clarendon Press.

- Gierke, Otto von. 1900. *Political Theories of the Middle Age*. Translated by Frederick Maitland. Cambridge: Cambridge University Press.
- Gordley, James. 2012. "Suárez and Natural Law." In *The Philosophy of Francisco Suárez*, edited by Benjamin Hill and Henrik Lagerlund, 209–29. Oxford: Oxford University Press.
- Jarlot, Georges. 1949. "Les idées politiques de Suarez et le pouvoir absolu." *Archives de philosophie* 18(1): 64–107.
- Jombart, Émile. 1932. "Le 'volontarisme' de la loi d'après Suarez." *Nouvelle revue de théologie* 59: 34–44.
- Moreau, Pierre-François. 1979. "Loi naturelle et ordre des choses chez Suarez." *Archives de philosophie* 42(2): 229–34.
- Pernoud, Mary Anne. 1972. "The Theory of the *Potentia Dei* according to Aquinas, Scotus and Ockham." *Antonianum* 47: 69–95.
- Pink, Thomas. 2012. "Action and Freedom in Suárez's Ethics." In *Interpreting Suárez: Critical Essays*, edited by Daniel Schwarz, 115–41. Cambridge: Cambridge University Press.
- Rommen, Heinrich. 1926. *Die Staatslehre des Franz Suarez S. J.* Mönchengladbach: Volksvereins.
- Skinner, Quentin. 1978. *The Foundations of Modern Political Thought. The Age of Reformation*. Vol. 2. Cambridge: Cambridge University Press.
- St. Leger, James. 1962. *The "Etiam si Daremus" of Hugo Grotius: A Study in the Origins of International Law*. Rome: Angelicum.
- Suñer, Pedro. 1974. "El Teocentrismo de la ley natural." In Francisco Suárez. *De legibus*, edited by Luciano Pereña et al., vol. 3; *De legibus (II 1–12): De lege naturali*, edited by Luciano Pereña and Vidal Abril, xxxviii–lv. Madrid: Consejo Superior de Investigaciones Científicas.
- Tierney, Brian. 2014. *Liberty and Law: The Idea of Permissive Natural Law, 1100–1800*. Washington: Catholic University of America Press.
- Treloar, John L. 1991. "Moral Virtue and the Demise of Prudence in the Thought of Francis Suárez." *American Catholic Philosophical Quarterly* 65(3): 387–405.
- Villey, Michel. 1975. *La formation de la pensée juridique moderne: Cours d'histoire de la philosophie du droit*. 4th ed. Paris: Montchrestien.
- Welzel, Hans. 1980. *Naturrecht und materiale Gerechtigkeit*. 4th ed. Göttingen: Vandenhoeck & Ruprecht.
- Westerman, Pauline C. 1998. *The Disintegration of Natural Law Theory: Aquinas to Finnis*. Leiden: Brill.
- Wilenius, Reijo. 1963. *The Social and Political Theory of Francisco Suárez*. Helsinki: Akaateeminen Kirjakauppa.

# Natural Right and Worldly Emotions of Love: Descartes, Molière, Spinoza

Both the classical and the scholastic natural right theories underwent a series of profound changes during the seventeenth century. The most important novel features can be captured easily if we compare the concise formulation of the classical concept in Cicero's *Laws* with the new concept of natural right that appeared in the seventeenth century. Cicero poses a question that proved to be crucial for the seventeenth century: "what has nature given to humans [?]".

In his answer, he creates a conceptual web containing law, nature and reason/rationality (*ratio*), ascribing to the latter independent commanding force both in man and the universe.

[L]aw is the highest reason (*ratio summa*), rooted in nature, which commands things that must be done and prohibits the opposite. When this same reason is secured and established in the human mind, it is law. [L]aw is judgment, the effect of which is such as to order people to behave rightly and forbid them to do wrong; [...] law is a power of nature, it is the mind and reason of the prudent man, it distinguishes justice and injustice. [...] I will seek the roots of justice in nature, under whose leadership our entire discussion must unfold. (Cicero 1999, 111–2)

The idea of the reason-based ethical community between men and the gods or the unique God is another peculiar feature that accompanied natural right theories until the seventeenth century. In Cicero it reads as follows:

[...] since there is nothing better than reason, and it is found both in humans and in god, reason forms the first bond between human and god. And those who share reason also share right reason; and since that is law, we humans must be considered to be closely allied to gods by law. (Cicero 1999, 113)

In the seventeenth century there were two parallel ways of philosophical-theological reflection, the main goals of which were very similar albeit they hardly took cognizance of each other. The one attempted to reformulate the theory of natural right from the perspective of the Protestant or Catholic reformation: Suarez, Grotius, Althusius and Pufendorf are the main representatives of this direction. The other intended to create a universally valid philosophical ethics without a traditional theology-based foundation. Its novel concept of a foundation was the new idea of the human being based on the mathematical-mechanical sciences.

In the works of the thinkers of the latter direction, in Descartes for example, we can hardly find explicit positive references to the natural right tradition. In Hobbes and Spinoza, who integrate in the thematic spectrum of the new philosophy the theory of society and also certain aspects of theology, one can only find sharply critical passages that provoked the reactions of Pufendorf from the other side. Within Leibniz's monumental *œuvre* that aims at the unification of traditional and new philosophy, theology and jurisprudence, the natural right also found a place among the provisional titles: *Elementa juris naturalis*<sup>1</sup> Unfortunately, however, this was only the title of fragments, excerpts, projects, not that of completed works.

In these projects, one of Leibniz's main opponents was Grotius, a member of the first group of thinkers, who set the stage for the subsequent treatments of natural right. In the theoretical introduction of his *De jure belli ac pacis* he reformulates Cicero's views as follows:

For by reason that Man above all other Creatures is endued not only with this Social Faculty of which we have spoken, but likewise with Judgment to discern Things pleasant or hurtful, and those not only present but future, and such as may prove to be so in their Consequences; it must therefore be agreeable to human Nature, that according to the Measure of our Understanding we should in these Things follow the Dictates of a right and sound Judgment, and not be corrupted either by Fear, or the Allurements of present Pleasure, nor be carried away violently by blind Passion. And whatsoever is contrary to such a Judgment is likewise understood to be contrary to Natural Right, that is, the Laws of our Nature. (Grotius 2005, 87)

The main theses of this passage would become vehemently disputed topics in the writings of subsequent philosophers, mostly without making clear the provenience of the issues. The basic problem was whether the human being is to be construed as a social being by nature or as a solitary individual who establishes societies by coercion in the way the atoms of the Epicurean connect for a while and disconnect afterwards due to the varying pressure of the neighbouring things.

---

1 Cf. the excellent collection edited by H. Busche, and the further literature referred to by the editor: Leibniz 2003.

Another topic from the passage of Grotius that would become important in the seventeenth and eighteenth centuries is the question by which faculty the human being distinguishes between what really good and really bad is with respect to ethical decisions, and if it is possible for her to resist the ever-changing external influences that take the form of various passions. According to Grotius and the natural right theoreticians, the distinction is made by reason or intellect that they understood to function well notwithstanding their natural and supernatural limitations; the latter was considered the consequence of original sin, while the former was caused by the influence of the passions, the traditional impediment of a reasonable life.

The third main topic maintained by Grotius is the insistence on the *obligation* character of natural *right*: one *ought to* use one's capacities to make adequate judgements concerning good and evil, and one *ought to* keep to one's judgements under the leadership of nature, *natura duce*, as Cicero had formulated. It is this complex construction that the best-known sentence of Grotius's whole work powerfully underlines:

And indeed, all we have now said would take place, though we should even grant, what without the greatest Wickedness cannot be granted, that there is no God, or that he takes no Care of human Affairs. (Grotius 2005, 89)

From our perspective, this implies first of all that "nature" as the proper abode for the functioning of reason or intellect is on the way to its later role of an independent authority. As nature founded the community of gods and men in Hellenistic philosophy, so it does in Grotius's natural right in the form of a particular kind of community between God and the *essence* of man:

And even the Law of Nature itself, [...] though it flows from the internal Principles of Man, may notwithstanding be justly ascribed to God, because it was his Pleasure that these Principles should be in us. (Grotius 2005, 91)

This sentence can be interpreted both from a humanistic and a theistic perspective, similarly to Descartes' double foundation of his philosophy in *God* and the *thinking I*. One of the main novelties in Descartes' philosophy consisted precisely in this double foundation that implied further the bifurcation of the concept of nature abbreviated as "God or Nature" both in Descartes and Spinoza.

This development of the redoubled foundation rendered the interpretive possibilities of the right of *nature* considerably complex. For to begin with, no classical theory of natural right can be expected where "nature" is not conceived as inseparable from reason as the godly ruler of cosmos and of human behaviour.

In this sense, it was a disquieting novelty in Descartes' *Meditations on First Philosophy* that all of a sudden the reader had to face the following statement:

For since I had apparently natural impulses towards many things which reason told me to avoid, I reckoned that a great deal of confidence should not be placed in what I was taught by nature. (CSM II 5/AT VII 77)

This is a striking assertion, indeed. If this were all what Descartes had to tell us about nature, it would be tantamount to discarding the whole tradition of natural right with the paradigmatic infallibility of nature in the background. This does not happen, however, for in the next step he identifies with “[my] nature or [my] essence” the most important result of the earlier meditations, the concept of *res cogitans* as separated from the body: “absolutely nothing else belongs to my nature or essence except that I am a thinking thing [...]” (CSM II 54/AT VII 78)

This identification is a consequent move on the way to the final statement about nature in Meditation 6 that establishes *in a new manner* a direct relationship and community between man and God.

Indeed, there is no doubt that everything that I am taught by nature contains some truth. For if nature is considered in its general aspect, then I understand by the term nothing other than God himself or the ordered system of created things established by God. And by my own nature in particular I understand nothing other than the totality of things bestowed on me by God. (CSM II 56/AT VII 80)

In the final analysis, therefore, Descartes, similarly to Grotius, traces nature back to God, even if this does not imply a unanimously positive evaluation of nature’s influence on us. Behind this ambiguity, we can suspect the different determination of the relationship between intellect and will in Grotius and Descartes. Although Grotius mentions the will, he subordinates it to the intellect. In contrast with this, although “officially” Descartes also subordinates the will to the intellect, he provides the will with the alarming licence to free itself from reason or intellect as a prefiguration of the merely “negative freedom” that would eventually play a distinctive role within a radically different theory of natural right. Once a thinker has made the decisive step to interpret “the teaching of nature” as an authority in which no full confidence is allowed to be placed, the way was open to a concept of natural right that relies on “nature” cut off from its foundational and governing reason. Such are – at least partially – the theories of Hobbes and Spinoza.

Nevertheless, Descartes remains much closer to the traditional concept of natural right than Hobbes. He enumerates a detailed list of the particular *truthful* teachings of nature. I quote only one of them, the one that belongs to the narrow circle of nature’s teachings that “God has bestowed on me as a combination of mind and body”, namely: “to avoid what induces a feeling of pain and to seek out what induces feelings of pleasure” (CSM II 57/AT VII 82).

This particular teaching is especially useful for us now because it can be connected directly to our above quotation from Grotius. The only slight difference is that Descartes does not call by name the cognitive faculty that helps us separate the feelings of pain to be avoided from those that must be endured for ethical or other reasons, respectively the feelings of pleasure to be avoided from those that can be approved. Descartes instructs us about how to carry out this separation partly in his last work, *The Passions of the Soul*, partly in his correspondence with Elisabeth of Bohemia.



*The Passions of the Soul* has a unique significance concerning the intricate relationship between natural right and the parallel philosophical undertaking of a reason-based ethics. To extricate this issue, we must first attentively read the famous closing sentence of the Letter to Picot usually published as a Preface to the treatise. This text is a good example of how an incautious interpreter can become entangled in the trap of relying too much on today's meanings of the words Descartes used.

[M]y intention was to explain the passions only as a natural philosopher [*en Physicien*], and not as a rhetorician or even as a moral philosopher. (CSM I 327/AT XI 326)

The Greek *physis* behind the French *en Physicien* refers to nature in the same way as the Latin *natura* and the French *nature*. Thus, far from being a univoque sign, the term *en Physicien* evokes as much the ambiguity within *nature* the “unreliable natural impulse”, and the “truthful God-given teaching” as the expressions *natura docet* or *institution de la nature* that appear several times in the *Meditations* and *The Passions of the Soul* respectively. The expression *Dieu ou la nature* figures more often in *The Passions of the Soul* than *Deus sive natura* in the “pantheist” Spinoza's *Ethics*. Therefore, the definition of nature that we have read in the version of Meditation 6 remains valid in the treatise on the passions as well; *en Physicien* must not be taken in the strong sense of today's natural sciences staying on the ground of *physicalism* as an at least methodological materialism of the finite sensual human being cut off from the rational sphere of the cosmic-divine. This remains true even if Descartes endorses top level brain science and physiology in more than the half of the articles in *The Passions of the soul*. That is because Descartes' physicalism must in no way be separated from the background idea of *creatio continua* maintaining that the status of the created world, nature at any given moment, depends on God's continuous creation. Nature as the result of the *continuous creation* is not the sum total of everything that *remained* from a primordial divine act, such as the ray of a light coming from a distant sun that can even have disappeared meanwhile. Descartes' nature is the sum total of everything that makes man an individual living being consisting of body and mind based on the always contemporaneous divine creation. And his being the image and likeness of God must be considered a fundamental distinctive difference between man and other beings insofar as man has a distinctive access to truth through his nature provided to him by the *verax* God.

This has serious consequences with respect to the allegedly other intention of Descartes when composing the treatise on the passions: the rejection of moral philosophy as such. For it is obvious that the text is full of sentences that must be understood as at least “soft” imperatives. Discarding the moralizing attitude toward the passions must, therefore, not be understood as discarding ethics in general. Descartes' real intention is to replace revelation-based moral theology with his own philosophical ethics as the basis of normative ethical statements. Thus it must have appeared evident for Descartes that the foundational theory of ethics in ancient thought before Christian moral theology took shape – natural right-based ethics – must be given a prominent role, even if this insight does not necessarily



seem evident for twenty-first-century interpreters. This prominent role appears first in § 79, the definition of love.

Love is an emotion of the soul caused by the motion of the spirits, which impels the soul to join itself willingly to objects that appear to be agreeable to it. [...] I say that these motions are caused by the spirits not only in order to distinguish love and hatred (which are passions and depend on the body) from judgement that also bring the soul to join itself willingly to things it deems good, but also to distinguish them from the emotions that these judgements produce in the soul. (CSM I 350/AT XI 388)

A key element in this passage is the expression “willingly” (*de volonté*) that Descartes interprets in a separate article (§ 80):

Moreover, in using the word ‘willingly’ I am not speaking of desire which is a completely separate passion relating to the future. I mean rather the assent by which we consider ourselves henceforth as joined with what we love in such a manner that we imagine a whole, of which we take ourselves to be only one part, and the thing loved to be the other. (CSM I 350/AT VII 388)

In my view, this auxiliary passage integrates the definition of love into Descartes’ system in which the *instant*, the moment must be viewed as one of the central concepts. The assent given willingly does not extend beyond the moment.

No doubt, this conceptual construction does not resemble our everyday idea of love. But Descartes’ aim is far more than merely interpreting our everyday concepts. He points out the systematic place where the basic factors determining the normative-evaluative attitude become integrated into the descriptive line of his thought. When judging a thing to be *agreeable* to us we posit by an act of will the existence of a whole that consists of the loving I and the thing beloved. By this act, the will sets up a new reflexion-based centre of value, as opposed to the self-enforcing valuing of the loving I’s own person, which is based on a *natural impulse*. The new centre is the whole the two parts of which are the respective I and the thing beloved *whatever this latter may be*. Already looking forward to Spinoza, I would particularly emphasize that the thing constituting the second half of the whole can be whatever the lover deems to be agreeable to her. Everyday people are prone to yield to the first superficial appearance that proves soon to be an illusion. Spinoza’s treatment of *vana religio* in the preface to the TTP is based on this feature of the finite existence.

For Descartes, the basic examples of the worthiest objects of the wise person’s love are her prince, country, and town – worldly beings that are entitled to claim for themselves the *devotional* love that we owe first and foremost to God.<sup>2</sup> At the same time, shockingly,

---

2 On the basis of this idea of devotion together with the second part of the *Discourse on Method* and some of the letters to princess Elisabeth and Chanut, it is not difficult to reconstruct Descartes’ monarchist

at first glance, Descartes takes someone attempting to rape someone else to be an example of love. There can be no doubt that in this case of love, one conceives the nature of both herself and the “beloved” in the most obscure manner possible, whereas in the examples of virtuous love the lover has clear and distinct perceptions of the nature of both herself and the objects of her love. The aggressive lover has but a highly obscure concept of what and how can be agreeable to her nature – and also of her own nature itself; either she considers it separated from the source of her existence, i.e. God as the highest good and the most perfect being, or she conceives of the good itself mistakenly. In contrast with this example, those who see in their children their other selves have much clearer concepts of their own nature, the nature of the thing beloved, and the divine goodness and perfection that is realized in the world continually.<sup>3</sup> Thus, behind the seemingly unified, neutral and descriptive concept of love in Descartes, descriptive and normative aspects of love can be separated on the basis of the differentiation of the modes of cognition that play a decisive role in the establishment of, and shaping the voluntary unification of the I and the thing beloved. It is obscure and unreliable in the one type of love, whereas clear and distinct in the other.

Thus, although Descartes does not explicitly refer to Grotius, we can see now how he agrees with him and the whole natural right tradition concerning the obligation of man to fight against the passions that obscure the eyes of the mind. Descartes takes over both the conceptual framework and the principal weapon of man, i.e. the evident insights of reason (Cf. *Passions* §§ 40 sqq).

As the second step, Descartes applies the traditional metaphor of the love of our other selves to the love of our children, and the interpretation of the relationship we call love in everyday parlance. However, as mentioned above, he refuses to consider this passion love in the proper sense. Instead, he subordinates it to desire as the sixth of the primary passions he acknowledges. It is instructive to see the ambiguity in his talks about nature and perfection in connection with this “highest desire”.

But the principal attraction comes from the perfections we imagine in a person who we think capable of becoming a second self. For nature has established a difference of sex in human beings, as in animals lacking reason, and with this she has also implanted certain impressions in the brain which bring it about that at a certain age and time we regard ourselves as deficient – as forming only one half of a whole, whose other half must be a person of the opposite sex. In this way, nature represents, in a confused manner the acquisition of this other half as the greatest of all goods imaginable. Although we see many persons of the opposite sex, yet we do not desire many at any one time, since *nature* does not make us **imagine** that we need more than one other half. But when we observe something in one of them which is more attractive than anything that we observe at that

---

political theology and a rather conservative theory of natural right that almost prefigures Burke's critique of the French revolution.

3 The above examples are taken from § 82.

moment in the others, this determines our soul to feel towards that one alone all the inclination which *nature* gives it to pursue the good which it represents as the greatest we could possibly possess. The name ‘love’ is applied more often to the inclination or desire which arises in this way from attraction than to the passion of love described previously. (CSM I 360/AT XI 395)

In this passage, the ambiguity has redoubled. The first ambiguity inheres in the concept of nature that reappears as an important quasi-agent in this concept of love-desire. For, to recall the result of the above analyses, “nature” signifies two things: first, the basic drives that function similarly in men and animals; this is the descriptive meaning of nature. Second, in human beings, it signifies the strivings *based on reason* that goes beyond the imagination, the type of cognition responsible for obscurity. This is precisely what constitutes the foundation of the *normative* teaching of nature in the natural right tradition. The normative teaching of reason-based nature prescribes *perfection* as the appropriate object of the *desire* called “love”. But this is in disagreement, even contrast with the way nature in the descriptive sense functions in the beings of animal nature. For evidently, Descartes’ view that everyone is (or to put it in a way that unveils the hidden soft imperative: must be) attracted by *one unique person* as the only appropriate other half is indifferent concerning perfection or imperfection. The archetype of this view is the speech of Aristophanes in Plato’s *Symposium* (cf. Plato 2011 189a sqq.). Descartes’ own version of it is related to his mechanical laws of movement and their descriptive natural-scientific explanation.<sup>4</sup> The

---

4 It is here that Descartes’ distinctive claim that *love is measurable* begins playing a significant role. As in the case of colliding bodies, where the greater body determines the joint movement of the two making up a new unit, if we have “colliding” lovers, the interests of the “greater” will determine the actions to be assented to by the “smaller”. Descartes’ tripartite division of love represents the three basic possible relations of magnitude between the lovers: “We may [...] distinguish kinds of love according to the esteem which we have for the object we love, as compared with ourselves. For when we have less esteem for it than for ourselves, we have only a simple affection for it; when we esteem it equally with ourselves, that is called ‘friendship’; and when we have more esteem for it, our passion may be called ‘devotion.’” Descartes claims that this division explains the *essence* of love rather than its *effects*, as the traditional distinction between concupiscent and benevolent love does (cf. § 81). The essence of love is what provides us with value-based action-readiness: in all types of love, “we consider ourselves as joined and united to the thing loved, and *so we are always ready to abandon the lesser part* of the whole that we compose with it so as to preserve the other part” (emphasis added). The interest of the “greater” ought to be valued more than the interest of the “lesser”: “pure description” is claimed to establish *moral evaluation*. The physical, *descriptive* law is transformed into a *normative* law establishing moral values and making us act according to these values. The first and third kinds of love complement each other: in every relationship between *unequal* parties the lesser loves “devotionally,” whereas the greater loves “affectionately”: “In the case of simple affection this results in our always preferring ourselves to the object of our love. (Ibid); [...] In the case of devotion [...] we prefer the thing loved so strongly that we are not afraid to die in order to preserve it.” (Ibid.) Thus, becoming part of a love relation *ab ovo* establishes moral values, and loving the greatest possible object implies that it is endowed with the highest moral value. For Descartes, the most appropriate object of love is God, the main effect of such a love being that the lesser part of the whole – man – ought to feel obliged to sacrifice himself for the

decisive element of the explanation can be rightly called the “principle of discontinuity” in reverse analogy to the way Leibniz would famously talk about his “principle of continuity”. The person who attracts the lover even by an infinitesimally small amount more than the ones after her, acquires automatically the exclusive right to occupy the status of the unique completing other half.

Yet, similarly to the way in which Leibniz refused the mechanical application of the same principle, one can even more easily refuse the application of the principle to the passion of “love as desire” by way of a reference to the instinctively accepted principle of the “continuity of attractions”. Let us allow for a moment that the quantity of love-attractions can be measured. Now, for every set of attractive members of the other sex we will always find one who attracts at least a little bit more than the others. Still, this fact alone will hardly suffice to prove the normative claim to the subsequent exclusive acceptance of the love enkindled by this one attractive man or woman. Descartes should have continued to line up forceful arguments in favour of his thesis, which is no less than the mechanical philosophical equivalent of what can be translated in everyday parlance as the priority of monogamy established by nature. And what is even more, it is *reciprocal* monogamy that Descartes must have had in mind as nature’s obligation, and this seems to be even more difficult to be proven. The forceful arguments could have been based on the traditional moral theology going back to Thomas Aquinas; but Descartes declined to rely on it. Another candidate could precisely have been the normative concept of reason in the pre-Christian natural right – the one that would remain intact even if God did not exist or care for human affairs, as Grotius formulated it.

In all probability, the profound crisis of the contemporary theory of natural right rendered it impossible for Descartes to go as far as this implied on the usual way of natural right theoreticians. An essential aspect of this crisis can be excellently pointed out in an influential antagonist of Descartes the normative philosophical moralist: Molière, the sharp-eyed playwright.

That our everyday experience of and thinking about our individualistic human behaviour cannot be considered analogous to the mechanics of the simplest bodies so easily, nor can it be tamed so simply by invoking traditional natural right categories, as Descartes wished to do, is manifested in a well-known passage from Molière’s *Don Juan* that seems to be unexploited in this respect.

Don Juan: What! Would you have a man bind himself forever to the first beauty that captivates him, renounce the world for her, and never more have eyes for any other? T’is a fine thing indeed to pride oneself on the false honour of being faith-

---

greater part – God. “As for devotion, its principal object is undoubtedly the supreme Deity [...]” (Ibid.) Furthermore, God’s right to obligate one to a sacrifice is transferred to the sovereign, who has the right to obligate his subject to die. “But we may also have devotion for our sovereign, our country, our town [...] We have often seen examples of such devotion in those who have exposed themselves to certain death in defence of their sovereign or their city [...]” (Ibid.)

ful, to bury oneself forever in one passion, and from our very youth to be dead to all the other beauties that may meet our eyes! No, no. Constancy (*constance*) is fit only for fools; all the fair sex have the right (*ont droit*) to charm us, and the good fortune of one in being met with first ought not to rob the others of their just claims (*juste prétention*) on our hearts: [...] Even though I 'm bound to one fair charmer, the love I may feel for her does not compel my heart to do injustice to the others (*faire injustice aux autres*); I still have eyes to see the worth of all, and I pay to each the homage and the tribute that nature demands of us (*la nature nous oblige*). (Part 1, Scene 2: Molière 1908, 67–8)

In this monologue, Don Juan forcefully parodies both the juristic principle of the “first apprehension” (*prior apprehensio*), the Cartesian principle of the “the biggest charm wins all”, and the normative statement of reason in favour of reciprocal monogamy. He acknowledges the freshly discovered “natural right” of every being to do what follows from its nature; for if we generalize what he says concerning women we arrive to this point. Thus he discards both the theological and the natural right-ethical arguments in favour of monogamy. Love does not comply with reason in any of its forms, and every form of partnership must be determined by *passionate* love renouncing to reason, because this is what “*nature demands of us (la nature nous oblige)*”. It goes without saying that the term “nature” must be taken in its descriptive meaning here.

In Don Juan’s parody of the reason-based natural right, an a-rational “natural right” makes its provocative appearance that would come to its full-blown form in Spinoza’s TTP. According to this concept of a new natural right, what really governs nature is the principle of continual yielding to the continual attractions, and there is no argument that obligates us to distinguish a being with particular attraction and bind ourselves to it forever. Molière’s text is a provocative rejection of the above-quoted thesis of Descartes: “Although we see many persons of the opposite sex, yet we do not desire many at any one time, since *nature* does not make us imagine that we need more than one other half.”

The emphasised French expressions in Molière’s text make it quite evident that Don Juan employs the *termini technici* of the contemporary stoic natural right philosophy; this is what the text of the negative hero attempts to make ridiculous. It is in Spinoza’s work that this double relationship to the whole natural right problematic becomes *philosophically* explicit. He makes the first overall attempt to convince us of the redemptive force of morality that does not expect reward for virtuous deeds. His argumentation is a thoroughly humanistic way of: God exists for him but His existence is a principal instead of a personal one, and so He or It does not care for the human affairs.

The generalized concept of this a-rational natural law is presented *expressis verbis* in the TTP as follows:

By the right and established practice of *nature (jus et institutum naturae)* I mean nothing but the rules of the *nature* of each individual, according to which we conceive each thing to be *naturally* determined to existing and having effects in

a certain way. For example, fishes are determined *by nature* to swimming, and the large ones to eating the smaller. So it is by the supreme *right of nature* that fishes are masters of the water, and that the large ones eat the smaller.

For it's certain that *nature*, considered absolutely, has the supreme right to do everything it can, i.e. that the *right of nature* extends as far as its power does. For *the power of nature is the power of God* itself, and he has the *supreme right* over all things. But the universal power of the whole of nature is nothing but the power of all individuals together. From this it follows that each individual has a supreme right to do everything it can, or that the right of each thing extends as far as its determinate power does. Now the *supreme law of nature* is that each thing strives to persevere in its state as far as it can by its own power, and does this, not on account of anything else, but only of itself. From this it follows that each individual has the *supreme right* to do this, i.e. (as I have said) to exist and have effects as it is naturally determined to do.

Nor do we recognize here any difference between men and other individuals in nature, nor between men endowed with reason and those others who are ignorant of true reason, nor between fools and madmen, and those who are sensible and sane. For whatever each thing does according to the *laws of its nature*, it does with *supreme right*, because it acts as it has been determined to do according to nature, and *cannot do otherwise*.

So among men who are considered as living only under the rule of nature, one who does not yet know reason, or does not yet have a habit of virtue, has a *supreme right* to live according to the *laws of appetite* alone – just as much as one who guides his life according to the laws of reason. I.e., just as the wise man has the *supreme right* to do everything which *reason dictates*, or to live according to the *laws of reason*, so also the ignorant and weak-minded have the supreme right to do everything appetite urges, or to live according to the laws of appetite. (C 280 sq./G III 189 sq.)

Although nature as a whole is governed by a preeminent reason or intellect inaccessible for human beings, its evolutive dynamism (*potentia*) is a-rational from the point of view of the individual human being. Accordingly, God's essence can solely consist in its all-might-power (*potentia*) (cf. E1p34). This peculiar approach to nature, man, and their relationship to each other is expressed in the a-rationality of the natural right that our quotation brings to the fore again and again emphatically – so emphatically that it could not fail to evoke Pufendorf's likewise emphatically forwarded critical statement:

Here it is observable, that by the Term of Right he does not express any Law directive of an Action, but only the Power of acting and what may be done without Injury, and that therefore 'tis by no Means a fair Conclusion, that one ought necessarily to do all those things, which one has a Right of doing. Farther, as it is an improper Acceptation of the Term of *natural Law* to make it denote that according to which

every thing acts by a certain and determinate Manner, so is it likewise improper to apply the Name of Right to that Power and manner of acting which appears in irrational Beings: For he alone can be truly said to have a Right of acting, who acts on previous Reason and Deliberation.<sup>5</sup>

No doubt, Pufendorf is right from the point of view of traditional natural right theory. But it is important to notice that what Spinoza would like us to do is not to ignore *all* difference between rational behaviour and behaviour refusing every reasonable deliberation. For, on the one hand, what he wants us to do is to reject as inadequate the traditional natural right terminology – similarly to Descartes in this respect, who develops an ethical philosophy, but declines to use the terminology of moral theology. On the other hand, what Spinoza has to say about the superiority of *rational* behaviour over the irrationality of obscure *imagination* can well be connected to Grotius’s conviction that the difference between the modes of cognition absolutely matters. Human reason and intellect are not (Spinoza) or not essentially (Grotius) corrupted by the Fall, and thus they are capable of adequately cognizing nature – at least that respective part of nature that affects us most. In Spinoza’s E5p4cs all the intellects knowing nature adequately “constitute [together] God’s eternal and infinite intellect” (C 615/G II 306) – as if we read Cicero about the community between the divine and human beings with some Neoplatonic flavour. And indeed, the virtuous conduct of the Spinozean sage (*sapiens*) who possesses an adequately understanding mind is in no way inferior to the moral ideal determined by the stoic natural right. The Spinozean sage shares the most important features of the Stoic sage, even with respect to the sociability that follows from his “real nature”. First, he keeps a safe distance from the “crowd” – Spinoza’s famous multitude – but, second, he establishes the community of the free people who are on the way to wisdom, i.e. becoming sage, while at the same time he does everything within his power to integrate among the free people as many members of the multitude as possible. After all, if we search for the germs of this special type of thought that is liberal *and* communitarian at the same time, we can find one in Descartes’ concept of an exclusive friendship interpreted by the analogy of our only other self, other half. Real, enduring friendship can only be conceived between virtuous people, as the traditional Aristotelian-stoic idea has it. It seems as if Spinoza had developed this concept into a whole friendship-based community, similarly to the way he transformed Descartes’ metaphysical doctrine of the “one substance – one attribute” into his own metaphysics of the “one substance – infinitely many attributes”.

---

5 “*hoc loco per vocabulum juris non notari legem, juxta quam agendum sit, sed facultatem agendi, & quid citra injuriam agere quis possit: eoque non recte inferri, omnia ad quae agenda jus aliquis habet, necessario quoque eundem agere debere. Deinde uti impropria est acceptio vocabuli, legis naturalis, quando notat id, secundum quod unaquaeque res certa ac determinata ratione agit: Sic et improprie potentia & modus agendi, qui in creaturis ratione non praeditis conspicitur, juris nomine vocatur. Nam proprie illum dumtaxta jus ad agendum habere dicimus, qui praevia ratione agit.*” (My translation, Book 2, chapter 2. Pufendorf 1998, I, 117)



There is no singular thing in Nature that is more useful to man than a man who lives according to the guidance of reason. 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 IID2), and only to that extent must he always agree with the nature of the other man (by P35). Therefore, among singular things there is nothing more useful to man than a man, etc., q.e.d. (E4p35c1; C 563/G II 233)

The social ethics that predominates among the free men led by the *rational, normative* nature is described in terms of Roman stoic ethics.

Again, whatever we desire and do of which we are the cause insofar as we have the idea of God, or insofar as we know God, I relate to Religion (*religio*). The Desire to do good generated in us by our living according to the guidance of reason, I call Morality (*pietas*). The Desire by which a man who lives according to the guidance of reason is bound to join others to himself in friendship (*amicitia*), I call Being Honourable (*honestas*), and I call that honourable which men who live according to the guidance of reason praise; on the other hand, what is contrary to the formation of friendship, I call dishonourable (*turpe*). (E4p37s1; C 565–66/G II 235)

Yet, this beautiful line of thought reminding the reader of ancient stoic ideas was far from real life in Spinoza's age – as far as ever, probably. Spinoza was aware of the unreality, and so he wrote in the last line of the *Ethics* that “all things excellent are as difficult as they are rare”. (C 617 / G II 308) This was also his reason for looking for methods in order to accomplish at least a considerable part of the way to the intellectual love of God shared by everyone who adequately understands nature and so are united and constitute God's infinite intellect – as our above quotation put it.

At the same time, when reflecting upon the philosophical relevance of natural right we must keep in mind that Spinoza was not a learned jurist practitioner – like Grotius, Pufendorf or Leibniz. He was a self-made philosopher, representative of the second group of thinkers whom I referred to in my introduction, and who set for themselves as their goal to create a metaphysically warranted ethics. He was a newcomer who had the best chances to contribute to the establishment of what we can call with Kuhn a new paradigm of natural right thought.<sup>6</sup>

---

6 On the other hand, the thinker who had the most chance to connect metaphysical philosophy to jurisprudence was Leibniz. It seems to be a matter of course to integrate him into our line of thought. Happily, we have Patrick Riley's seminal articles about this topic in Leibniz's oeuvre, and I myself have also written about it in his footsteps, so I can refer to these articles now: Riley 1996, Boros 2007, Boros forthcoming.



# Bibliography

Passages of Spinoza's *Ethics* are cited with the usual abbreviation; E stands for Ethics which is followed by the number of the part and then: p – proposition, s – scholium, c – corollary, d – definition if it is immediately after the number of the part and demonstration in all other cases; E2p11d is this the demonstration of proposition 11 of the second part of the Ethics.

## Primary Sources

- AT Adam, Charles and Tannery, Paul (eds). 1996. *Oeuvres de Descartes*, 11 Vols. Paris: Vrin.
- C Spinoza, Benedictus de. 2016. *The Collected Works of Spinoza Vol. II*. Translated by Edwin Curley. Princeton, N.J.: Princeton University Press.
- CSM Cottingham, John, Stoothoff, Robert and Murdoch, Douglas (trans). 1985. *The Philosophical Writings of Descartes*. Vols. 1–2. Cambridge: Cambridge University Press.
- G Spinoza, Benedictus de. 1925. *Opera, im Auftrag der Heidelberger Akademie der Wissenschaften*. Edited by Carl Gebhardt. Heidelberg: C. Winters.

Cicero. 1999. *On the Commonwealth and On the Laws*. Translated by J. E. G. Zetzel. Cambridge: Cambridge University Press.

Grotius, Hugo. 2005. *The rights of war and peace*. Edited by Richard Tuck. Indianapolis: Liberty Fund.

Leibniz, G. W. 2003. *Frühe Schriften zum Naturrecht*. Edited by Hubertus Busche and Hans Zimmermann. Hamburg: Meiner.

Molière. 1908. *The Affected Misses, Don Juan, etc.* Transl. by C. H. Page. New York – London: G. P. Putnam's Sons.

Plato. 2011. "Symposium". In *Lysis, Symposium, Gorgias*, translated by Walter R. M. Lamb, 73–246. Cambridge, Mass: Harvard University Press.

Pufendorf. 1998. "De jure naturae et gentium." In *Gesammelte Werke*, IV/1. Berlin: Akademie Verlag.

## Secondary Sources

Boros, Gábor. 2007. "Leibniz on Love." In *The Concept of Love in 17th and 18th Century Philosophy*, edited by G. Boros, H. De Dijn, M. Moors, 79–94. Leuven–Budapest: Leuven University Press – Eötvös University Press.

Boros, Gábor. 2020. "A Political Theology 'of Doubtful Solidity'. Leo Strauss on Rousseau via Spinoza." In *Past and Present Political Theology: Expanding the Canon*, edited by D. Vanden Auweele and M. Vassányi. New York & London: Routledge.

Riley, Patrick T. 1996. *Leibniz' Universal Jurisprudence: Justice as the Charity of the Wise*. Harvard University Press.

# The Concept of Simple Esteem in *De jure naturae et gentium*

Recent interest in the concepts of human dignity and its close links to natural equality in seventeenth-century philosophy has largely been motivated by the predominance of the concept of dignity in Kantian moral philosophy and contemporary human rights discourse (e.g. Rosen 2012; Debes 2017). As a result, intellectual historians and philosophers have started to pay renewed attention to the concepts of dignity (*dignitas*) and natural equality in Pufendorf's philosophy. In particular, scholars have questioned the idea that we could read Pufendorf's employment of the concept of dignity as a precursor of Immanuel Kant's idea of human dignity as an end in itself (See Saastamoinen 2010; Haakonssen 2010; Zurbuchen 2019). However, given that Pufendorf was the first early modern philosopher who paid extensive attention to the concept of esteem and used it lengthily in his works (Hruscka 2000), it is striking how little attention Pufendorf's treatment of the concept of esteem (*existimatio*) has received so far.<sup>1</sup>

This paper provides the first attempt to systematize the role of simple esteem (*existimatio simplex*) in Pufendorf's magnum opus *De jure naturae et gentium*, 1672 (hereafter *De jure*).<sup>2</sup> My analysis is restricted to *De jure*, which represents the most comprehensive

- 
- 1 Some of the textual material of this essay is based on the analysis presented in Haara 2018, chapter 4, which focuses on the psychological desire for esteem in Pufendorf's theory of sociability.
  - 2 The original Latin version of *De jure* is cited first, separated by a forward slash from the translation. I am referring to *De jure* by the organizational scheme that Pufendorf established. The standard scholarly edition of *De jure* is Pufendorf, Samuel. 2002. *De jure naturae et gentium*. Samuel Pufendorf, *Gesammelte Werke*, Band 4, edited by Frank Böhling. Berlin: Akademie Verlag [= JNG]. Whenever possible I use Michael Seidler's translation, Pufendorf, Samuel. 1994. *The Political Writings of Samuel Pufendorf*, translated by Michael Seidler and edited by Craig L. Carr. New York: Oxford University Press [= PWSP]. In the passages omitted by Seidler's translation, I employ Pufendorf, Samuel. 1934. *Of the Law of Nature and Nations*, translated by C. H. Oldfather and W. A. Oldfather. Oxford: Clarendon Press [= LNNO].

and developed account of Pufendorf's moral and political thought.<sup>3</sup> The purpose of the examination of Pufendorf's usages of the concept of simple esteem is to offer a better understanding of how he conceptualized the basic moral value of individuals. I maintain that Pufendorf's treatment of simple esteem should be scrutinised in its own right, rather than confused with his usage of the concept of dignity. The central difference between dignity and simple esteem is that the latter signifies humankind's superior value in relation to other animals, whereas the former denotes the most basic comparative value among human beings. In what follows I shall consistently translate the term *existimatio simplex* as "simple esteem" because Pufendorf employs *existimatio* as a moral philosophical concept that has a specific purpose in his moral and political philosophy.<sup>4</sup> On my reading, *existimatio* signifies people's moral value in relation to others in a similar way as price denotes the comparative moral value of things.

The secondary aim of this paper is to offer historical perspective to the contemporary debates on social esteem by shedding light on the overlooked aspects of the conceptual history of esteem. In current political theory, it has often been noted that social distinctions in contemporary civil societies are seemingly more open and flexible than in premodern societies. In the theories of recognition in particular, the concept of esteem has been used in the analysis of the normative character of social distinctions in contemporary liberal democracies (e.g. McBride 2013, 72–102). It has been argued that one of the central features of modern capitalistic societies is that an individual's esteem is markedly decided by his or her accomplishments as a "productive citizen" (Honneth and Fraser 2003, 143). This idea bears similarities to Pufendorf's theory, in which the gradation of simple esteem is grounded on one's performance as a sociable citizen, not on the traditional feudal idea of estate-based status hierarchy. It is impossible and unproductive to try to establish a direct link between Pufendorf's notion of simple esteem and how the concept of esteem is used in present-day political philosophy. However, I hope that this essay invites modern scholars interested in the multifaceted concept of esteem to reflect on current theoretical ideas on social esteem by comparing them with those offered by Pufendorf.

This paper comprises three sections. The first section explicates the concept of simple esteem in Pufendorfan moral vocabulary. The second section deals with the different functions of simple esteem in a state of nature and in a civil state. The third section explores the natural right of simple esteem in Pufendorf's theory of natural rights.

---

3 Pufendorf also dealt with *existimatio* in his earlier *Elementa Jurisprudentia universalis* (1660) and in the dissertation *De existimatione* (1667). These earlier writings fall outside the scope of the present paper.

4 In early eighteenth-century translations and modern translations of *De jure*, *existimatio* is translated as either "esteem" or "reputation" depending on the context. In the first German translation this term was translated as "die Achtung". For the problematic of translating *existimatio* into vernacular languages, see Hruschka 2000, 91.

# Simple Esteem as a Moral Quantity

In order to make a thorough interpretation of Pufendorf's treatment of simple esteem in *De jure*, it is imperative to start by introducing his renowned demarcation between moral entities and natural entities, which provides a conceptual framework for his moral theory. In Pufendorffian moral science, the world is divided into two kind of entities, physical entities (*entia physica*) and moral entities (*entia moralia*). The objects of the natural world belong to the domain of physical entities. Moral entities are imposed on an already existing physical world by intelligent beings. The most fundamental moral entities are assigned into the world by God, jointly with his creative act. However, people also have the ability to produce additional moral entities that regulate moral and social life. Moral entities cover a wide range of different norms and institutions that produce moral effects and direct human actions. Pufendorf explains that moral entities are "not self-subsistent but depend on substances and their motions, which they only affect in a certain manner" (*JNG* 1.1.3/*PWSP* 100). While moral entities rely ontologically on self-subsistent physical entities, they do not function on the same level as the physical realm. Moral entities lack physical and causal power and "their operative power consist not in directly producing some physical motion or change in things by means of their internal efficacy" (*JNG* 1.1.4/*PWSP* 100–101). The purpose of moral entities is to create order and harmony in human life by restricting the freedom of man's voluntary actions (*JNG* 1.1.5).

Pufendorf's moral ontology includes four kinds of moral entities: 1) states and conditions, 2) substances (such as moral persons), 3) moral qualities (affective modes) and 4) moral quantities (estimative modes). This paper focuses on esteem (*existimatio*), which fits in the class of moral quantities in Pufendorf's moral vocabulary. The features of the moral world, such as states, persons and things, are estimated in terms of their moral quantity. Pufendorf notably distinguishes between the moral quantity of things, labelled price (*pretium*), and the moral quantity of persons, called esteem (*existimatio*) (*JNG* 1.1.22). This differentiates Pufendorffian terminology from that of Thomas Hobbes's earlier equation of the price of things and the price of human beings in *Leviathan*. According to Hobbes, "(t)he Value, or worth of a man, is of all other things, his Price" (Hobbes 1991, 63). It is plausible that Pufendorf aimed to distinguish his theory from that of the controversial Hobbes by stressing the terminological difference between the value of things and the value of persons. Nevertheless, we should not exaggerate the actual differences between Hobbes's and Pufendorf's conceptualizations of the value of persons. Despite Pufendorf's demarcation between *pretium* and *existimatio*, he does not ground the esteem of persons on an inalienable respect for the person's metaphysical status as a human being.<sup>5</sup> Rather, in a similar way as Hobbes, he is concerned about the conditions of political stability and

---

5 I think that Stephen Darwall (2012, 224–5) makes this mistake in his otherwise interesting analysis on the interpersonal character of Pufendorf's theory of sociability.

sociability. Most importantly, varying artificial moral estimations are necessary after the founding of societies:

And as prices are set upon things chiefly to the end that in their exchange or transfer from one to another they can be correctly compared with each other, so also esteem serves to that end, that by, it, as by value, men can be compared with one another, so that a decorous order can be set up in the case they happen to be united, when it was found that a general equality among mankind could not be conveniently preserved (*JNG* 8.4.1/*LNNO* 1229–330).

This passage implies that the purpose of the imposition of price and esteem is the same, that is the establishment of peaceful social order and harmony. It should be noted that Pufendorf's above usage of the phrase "a general equality of mankind" does not refer to the notion of natural equality, but rather to the idea that people cannot establish functioning societies without imposing social distinctions. Despite the great variety of people's comparative social standings, all people are obviously naturally equal. Pufendorf thus notes that social interaction is not possible with an individual "by whom one is not esteemed at least a man". Therefore, "everyone must esteem and treat other men as his natural equals, or as men in the same sense as he" (*JNG* 3.2.1/*PWSP* 159).

Because of the distinction between nature and morality (between natural entities and moral entities), Pufendorf maintains that moral esteem cannot be directly inferred from nature. Thus, social distinctions are not grounded on the natural features of individuals. For instance, the sovereign's higher esteem is based on contract or consent, rather than on their natural ability to rule others (*JNG* 7.2.1). Pufendorf explains this as follows:

For it is evident that in common life persons and things are valued not only according to their extension as physical substances, in the degree of their motion as physical qualities – inasmuch as these are considered to spring from natural principles – but also according to another kind of non-physical and non-mathematical quantity arising from imposition and the determination of rational power. (*JNG* 1.1.22/*PWSP* 108)

The premise that moral effects cannot be straightforwardly deduced from physical entities structures Pufendorf's demonstration of moral norms and institutions throughout *De jure*. In his sociological observations, however, Pufendorf acknowledges that people erroneously tend to draw moral conclusions directly from physical and natural features of human nature. Against this kind of widely held belief, Pufendorf maintains that cleverness or the beauty of the body ought not to ground a person's moral esteem. For instance, there is no such a thing as natural nobility (*JNG* 8.4.25). Moreover, Pufendorf points out that the imposition of moral entities, such as a status of a noble, does not produce enduring moral character. Therefore, "if a commoner becomes a noble, he acquires merely new rights, but his substance and physical qualities are changed not one whit. If a noble be expelled from

his order, he loses merely his rights; all of his natural endowments remain unimpaired.” (JNG 1.1.23/LNNO 21).

Against Aristotle’s comments on the lack of certainty in moral affairs, Pufendorf claims that moral science (*scientia*) can achieve moral certainty, just as physical sciences achieve physical certainty (JNG 1.2.3). At the same time, however, he acknowledges that the method of obtaining demonstrative moral knowledge in moral science is different from that of physical sciences. This holds particularly in the measurement of moral quantities, such as esteem. The value of physical quantities can be precisely equated and measured. In turn, moral quantities arise by imposition and rely on the judgements of free persons, “whose judgement and pleasure is in no way subject to physical measurement; and so the quantity which they conceive and determine by their imposition, cannot be referred to a like measure, but retains the liberty and laxness of its origin” (JNG 1.2.10/LNNO 35). Because esteem is a non-physical and non-mathematical quantity, it is impossible to measure it as accurately as the quantity of physical objects which comprise the extension of physical substances. However, Pufendorf maintains that it is sufficient for the purposes of social and moral life “that persons, things and actions be roughly rated and compared”. Most significantly, demonstrative moral science can validate “how one person should be preferred to another”. While we cannot mathematically determine whether an individual “is two, three, or four times as good as that of another”, nevertheless it is possible to demonstrate “how one person should be preferred to another” (JNG 1.2.10/LNNO 35).<sup>6</sup>

In this context, Pufendorf is obviously speaking of a demonstrative moral science that requires intellectual abilities and extensive training. Citing Descartes’ *Les Passions de l’âme*, he associates just esteem and free will, arguing that “the only just cause for esteeming ourselves proceeds from the proper use of our free power of choice, and from the control we exercise over our will” (JNG 3.2.6/LNNO 339). However, while educated moral experts (like himself) can always correctly evaluate their own and other people’s moral value, there are various passages which paint a rather pessimistic picture of the rational abilities of the uneducated multitude to evaluate the moral value of persons (or things) in relation to natural law requirements. Most evidently, people tend to estimate their own value more highly than it deserves. This is mainly due to the self-regarding psychology of human nature. For Pufendorf, self-love is the most significant human inclination that leads an individual “indignant at any reflection upon his dullness and imprudence, and greatly incensed at such as boast of their own prudence as being superior to that of others” (JNG 3.2.2/LNNO 331–2). Like Hobbes, Pufendorf therefore underlines that people are naturally wired to care about their esteem, which is a central source of social and political conflicts. Unlike among other animals, whose desires are simple, among human beings there is “a struggle for honour and dignity”, which is the root cause of “envy, rivalry and hatred” (JNG 7.2.4/LNNO 969). The desire to distinguish oneself from others drives the formation of social distinctions. In addition to the unsocial effects of the desire for esteem, Pufendorf is also aware of the constructive role of this desire. Pufendorf and later John Locke recognized

---

6 Translation slightly modified.

that the desire for esteem is not merely a threat to moral and social life, it also encourages people to act sociably and internalize the content of the reciprocal duties of social morality (see Haara and Stuart-Buttle 2018).

It is crucially important to note that Pufendorf's usage of the concept of esteem should not be confounded with his understanding of man's higher value compared to other animals, that is, dignity (*dignitas*). Human dignity hinges on natural theological premises. Unlike other animals, God has given people the understanding and the will, which make them capable of complying with moral obligations and participating in social and moral life. The tension between man's antisocial and social tendencies is inherent in Pufendorf's demonstration of the norms and institutions required to maintain and cultivate sociability as natural law demands. Pufendorf denies that men are by nature political and social creatures, yet he simultaneously understands dignity as a kind man's potential capacity to cultivate sociability. In contrast to dignity (man's higher value compared to other animals) *existimatio* is "the value of persons in communal life (*in vita commune*) according to which they can be equated or compared with others, and ranked before or after them" (*JNG* 8.4.1/*PWSP* 253). Man's higher value compared to other animals is imposed by God. In turn, esteem is always imposed by human beings themselves. Whereas God has imposed dignity on the world conjointly with his creation of human beings, people impose esteem on already existing physical beings. Moreover, while dignity is a permanent and enduring feature of human life (unless God wishes to change human nature), esteem is always artificial and changes from one context to another in social life.

The analysis of this paper is limited to simple esteem. It is nevertheless important to note that Pufendorf further distinguishes between simple esteem (*existimatio simplex*) and intensive esteem (*existimatio intensive*). The most noteworthy difference between simple esteem and intensive esteem is that the former belongs to the class of natural rights that people have pre-socially, while the latter is only "imperfectly" obligating unless it is "supported by some express or tacit pact which is understood to reside in established custom" (*JNG* 8.4.7/*LNNO* 1252). Simple esteem signifies a person's basic moral standing in relation to others. It is possible for one's simple esteem to fall, yet it cannot rise above a certain level. Like intensive esteem, person's simple esteem is nevertheless relative to other people's simple esteem. Comparative standings do not necessarily have to be unequal (Neuhouser 2013, 214). For Pufendorf, the equal level of simple esteem is always comparative. While simple esteem cannot extend above a certain level, intensive esteem is supplementary to simple esteem. Intensive esteem calls for information about personal capacity and commendable deeds that elevate a person's value above other people who are equivalent in terms of their simple esteem (Haara 2018, 107). Intensive esteem is closely linked to the concept of honour, which is a form of social recognition or an external sign of intensive esteem. The practices of honour are founded on intensive esteem that has been elevated above the maximum level of simple esteem.



# Simple Esteem and Natural State and Civil State

The previous section placed the concept of simple esteem in Pufendorf's account of moral entities. In particular, it highlighted the importance of differentiating simple esteem from the concepts of dignity and intensive esteem. Put simply, simple esteem is a moral quantity that indicates people's most basic comparative value in relation to others. On the one hand, unlike intensive esteem, simple esteem cannot be elevated above a certain level, yet simple esteem also admits of degrees. On the other hand, in the same way as intensive esteem, the amount of simple esteem is contextually variable depending on social contexts. Pufendorf also notes that simple esteem and intensive esteem have different functions in a state of nature and a civil state. In a state of nature or natural liberty simple esteem means:

That a man deport himself and be regarded as a person with whom you may treat as a good man, who is also inclined to accommodate himself to the laws of human society, and therefore ready, to the best of his ability, to observe natural law in his relations with others. That is, as we say that a thing which is of some use in human life is of value, and that which is useless is of no value, so you may say that at least some value attaches to him with whom one can at any rate treat as with a social being. But if a man plainly shows his unfitness for society, as he does when he shamelessly spurns and tramples under foot the law of nature, and the duties owed to other men thereby, you may rightly conclude that he is a man of no value. (*JNG* 8.4.2/*LNNO* 1230)

The most striking statement in this passage is that disruptive and unruly behaviour will abolish a person's simple esteem thoroughly, that is, an antisocial being ought to be treated as "a man of no value". This claim is normative rather than descriptive. If someone is incapable of acting sociably, we should legitimately eliminate his simple esteem. Pufendorf seems to have in his mind the idea that it is the moral duty of honest people to strip the wicked of simple esteem because their actions have disastrous consequences for social life and the well-being of society. Thus, simple esteem is always conditional and can be eliminated. One's simple esteem may be significantly diminished or entirely lost by antisocial behaviour. Most strikingly, individuals who "carry on war not against their special enemies but against their fellow men" lose the "esteem by which we rank and grade other men" completely. Unless they leave off "that fierce and bloody manner of life, others must show no more mercy on them than on wolves and other savage animals" (*JNG* 8.4.5/*LNNO* 1232–3). Pufendorf is speaking here of pirates and highway robbers. What brands them so special is the way of life that makes them the common enemy of mankind. However, even pirates can recover their lost simple esteem if they change their way of life and make reparations to their victims (*JNG* 8.4.5).



In a natural state, the level of simple esteem hinges on one's external behaviour and its effects on sociability and established social customs. In a civil state, however, the level of simple esteem is tied to the decrees of civil laws.

Simple esteem among those who live within states consist in a man being held to be, at the least, a common but a full and entire member of a state, or, in other words, in his not having been declared by the laws and customs of the state to be a vicious member of the same, but in his being held to be one of its number and a unit in its composition. Furthermore, this esteem is wanting by reason either of a man's mere status or of a crime. (*JNG* 8.4.6/*LNNO* 1233)

Just as in the state of nature, individuals or groups may lose their simple esteem completely. For instance, it is possible that civil laws grant no civil standing to slaves who are regarded as property (*JNG* 8.4.6).<sup>7</sup> Likewise, serious crimes may extinguish a person's simple esteem, which is conditional on the laws of specific states (*JNG* 8.4.7). It is worth noting that, in analyzing men's value in civil societies, Pufendorf occasionally employs the term *dignitas civilis*, explaining that civil dignity is the "value which the state sets on a man in conferring upon him by way of honour, sovereignty, or office or the handling of any public undertaking, or even a name or title" (*JNG* 8.4.13/*LNNO* 1245–6). This clearly echoes Hobbes's definition of the civic value of man as dignity in chapter 10 of *Leviathan*.<sup>8</sup> This kind of usage originally has its roots in Roman law, in which *dignitas* referred to the respect due to one's rank or office.<sup>9</sup> It is important not to confound Pufendorf's usage of the term *dignitas* as a public manifestation of civil esteem and human *dignitas* with regard to other animals (Haara 2018, 102–3).

Some of Pufendorf's remarks imply that a citizen's simple esteem is wholly determined by the judgement of the sovereign. For instance, in taking a critical stance towards duelling, Pufendorf argues that "every citizen's esteem stands and falls by the judgement of his supreme sovereign and the laws, and virtue lies in being willing and obedient to the laws, without regard for the empty opinion of idle and mischievous men" (*JNG* 8.4.8/*LNNO* 1239). At the same time, however, the spheres of simple esteem also extend beyond the confines of the authority of the state. The following passage is worth quoting in detail:

Furthermore, it is also obvious that a man's simple natural esteem (*existimationem simplicem naturalem*) cannot be taken from him at the mere pleasure of his sover-

---

7 There is still a difference between a slave as property and other kinds of property. In his *De officio hominis et civis*, Pufendorf notes that because slaves are humans it is not permissible to treat them "like other property, which we may use, abuse and destroy at our pleasure" (Pufendorf 1991, 130).

8 According to Hobbes, "[t]he public worth of a man, which is the value set on him by the Commonwealth, is that which men commonly call DIGNITY. And this Value of him by the Commonwealth, is understood, by offices of Command, Judicature, publike Employment; or by Names and Titles, introduced for distinction of such Value." (Hobbes 1991, 63–4)

9 For an analysis of historical uses of dignity as a rank, see Waldron 2015, 13–46.

eign, save for some previous sin on his part, which, because of its own depravity, or the express sanction of the laws, is accompanied with infamy. For since such a power can in no wise make for the preservation or betterment of the state, that power is never understood to have been conferred upon sovereigns. Yet just as a sovereign can injure a citizen by banishing him from the state, so he can do the same depriving him of civil esteem, in so far, at least, as to deprive of those advantages which in a state attend unimpaired esteem. But his own natural esteem (*naturalis existimatio*) can no more be taken from a man than his honourable character. [...] And it appears certain that no citizen is obligated to sacrifice his own simple esteem to the state, or to undergo actual infamy for its advantage. For the evil deeds from which actual infamy arises, can neither be commanded by a state nor should they be performed by a citizen. (*JNG* 8.4.11/*LNNO* 1240)

This passage makes two interesting claims. First, Pufendorf argues that the sovereign cannot abolish the citizen's simple esteem by "mere pleasure". Pufendorf does not deny the possibility of the violation of citizen's natural or civil rights (*JNG* 7.8.4). Therefore, when the sovereign unjustly diminishes a citizen's simple esteem, he clearly acts outside the boundaries of his legitimate political authority. Second, Pufendorf uses the phrase "natural esteem", which is otherwise very rarely used in *De jure*, arguing that one's natural simple esteem cannot be eliminated by the sovereign. In this context, he uses the term natural simple esteem synonymously with the term simple esteem.

What does Pufendorf mean by the term natural esteem? I think that his definition of price in civil society offers a solution to this question. While exploring the role of price (the value of things), Pufendorf notes that in states the price of things is established in two ways: "One way is by the decree or law of those in authority, the other by the general valuation and judgement of man, with the further consent of those who are the parties to the bargain." (*JNG* 5.1.8/*LNNO* 686). There is a domain of "natural" value of things and persons which is not wholly under the sovereign's jurisdiction. This also marks one of the central differences between Pufendorf's and Hobbes's political theories. For Pufendorf, it is not the civil sovereign alone who can determine the actual level of the simple esteem of his citizens. I shall next explore how the idea that the simple esteem belonging naturally to persons is manifested in Pufendorf's theory of natural rights.

## Simple Esteem as a Natural Right

The notion of simple esteem as a natural right to which all people are entitled distinguishes simple esteem from intensive esteem. As I have already mentioned, the deserved share of intensive esteem requires either prior knowledge about someone's capacities or deeds or explicit pact or consent. Both of these premises are lacking in the case of simple esteem, which belongs to all of us naturally. Viewed from this perspective, it is thus understandable

that Pufendorf thinks that simple esteem (not intensive esteem) belongs to the domain of natural rights, explaining that nature has given certain rights to people, prior to human agreement, “such as life, body, limbs, chastity, simple esteem, and freedom” (*JNG* 3.1.1/*PWSP* 158). The inclusion of simple esteem in the list of natural rights involves an idea that all of us are naturally entailed to have simple esteem pre-socially, that is, prior any precedential human action. Simple esteem “belong[s] by nature to every man in equal share, and when no evil deed has preceded, all men must be judged to be equally honourable” (*JNG* 8.4.2/*LNNO* 1230).

The interpretation of simple esteem as a natural right obviously depends on the understanding of Pufendorf’s theory of natural rights as a whole. Scholars have viewed Pufendorf’s theory of natural rights through different perspectives.<sup>10</sup> In my view, an idea that natural rights are corollaries of the duties imposed by natural law best captures Pufendorf’s treatment of natural rights. Natural right to simple esteem should be understood as a means to live up to one’s obligations to the foundational principle of natural law, namely the cultivation of peaceable sociability. Pufendorf deduces natural right to simple esteem from the obligations that natural law imposes on people rather than obligations that others owe natural right holders. This interpretation highlights the aspects of Pufendorf’s natural rights theory that stand in contrast to the predominance of individual rights prevalent in contemporary liberal theories.

The supremacy of duties is exemplified in Pufendorf’s treatment of the natural right of simple esteem. Simple esteem should be predominantly taken as a natural duty which is a derivative of the principle of sociability. The natural right of simple esteem, like all other natural rights, is based on the will of God, who has commanded people to cultivate and maintain peaceable sociability. Note that Pufendorf’s moral theory centres on the principle of sociability. His basic intention is to demonstrate that, given human nature and the natural world, people are universally in need of the law of sociability (*JNG* 2.3.15). All moral and political duties are derived from the principle of sociability. People are obliged to maintain their simple esteem since it is a requisite means to fulfil the duties imposed on them by natural law, that is, to act as useful and sociable members of society (Saastamoinen 2008, 240). Because individuals need a certain level of simple esteem to promote sociability, others are prohibited from insulting or lowering their simple esteem for unjustified reasons. The reciprocal practices of simple esteem are a precondition of sociable life in *anonymous interactions*. The mutual trust necessary for sociability in large-scale societies

---

10 According to Knud Haakonssen (1998, 35–46), Pufendorf grounds natural rights on the obligations that other natural law imposes on people, rather than people’s obligations to others. In this picture, the duty to esteem and correlative right to simple esteem is a means of fulfilling the obligations imposed by natural law. In turn, Richard Tuck (1979, 160–1) holds that natural rights are based on the obligations that other people owe to the right holder. Kari Saastamoinen (2006, 241–248) maintains that there are two separate foundations for natural rights. Natural rights to natural right to simple esteem is corollary of the duties imposed by natural law, whereas the natural right of self-governance is based on the obligation to regard rights holders as their natural equals.

would be demolished without sufficient recognition of the simple esteem of others (Haara and Lahdenranta 2018, 22–3).

Simple esteem as a natural right includes certain normative commitments. People have a natural duty to esteem others as good men, that is, treat them as persons capable of performing the mutual duties of sociability. To retain one's simple esteem in a state of nature entails that

a person conduct himself as, and be considered as, someone with whom one can deal as with a good man, and one who is bent on accommodating himself to the laws of human sociability; someone, therefore who is prepared to observe the natural law towards other inasmuch as he can. (*JNG* 8.4.2/*PWSP* 253)

Because simple esteem belongs to all of us naturally, others are obliged to grant simple esteem to us also in the case when they do not have any knowledge about our personal traits and capabilities to promote sociability. The duty to attribute simple esteem to others becomes an essential precondition of social life because people are naturally inclined to react violently if their simple esteem is questioned by insults (*JNG* 7.1.4). To grant people the simple esteem they are entitled to removes “the just cause of hatred” (*JNG* 3.3.1). Without the duty of esteeming others as “good men”, sociable life is not possible at all. This is what Pufendorf has in mind when he states in *De officio hominis et civis*, in a rhetorical contrast to Hobbes, that we ought to regard someone who “is not our fellow citizen, or with whom we live in a state of nature, not indeed as our enemy, but as a friend we cannot wholly rely on” (Pufendorf 1991, 119).

However, simple esteem should not be interpreted as an inalienable individual right in Pufendorf's theory. Pufendorf's main interest in promoting a culture of duties rather than rights is most obviously illustrated in his reflection on the duties to cultivate one's mind (*cultura animi*), which are part of man's duties toward himself. All people are obliged to cultivate their minds in order to obtain simple esteem. If an agent's simple esteem is diminished due to prejudiced or false allegations, he or she must strive to regain simple esteem by all legitimate means.

The most necessary part of this cultivation is that in which a person learns how to set a just value on the things that especially stimulate the appetites of man. For upon this depends the question of how much one should concern himself about each one of them. Now that which is considered the most eminent among these, and is fitted to attract in a special way men of more noble mould, is reputation for worth and glory from which are born honour and fame. Regarding this mind should be so shaped that it will endeavour in every way to procure a simple esteem, or the reputation of being a good man; and this results as a rule from the observance of the natural law and its duty, while a disregard of it can furnish occasion for many ills to us. But when it is not within our power to repulse calumnies and

false opinions about us, the rectitude of our conscience will be our solace, and the fact that our innocence is established in the sight of God. (*JNG* 2.4.9/*LNNO* 242)<sup>11</sup>

This passage makes three noteworthy claims. First, the duty to learn to truthfully measure the esteem of oneself and others is the most important part of *cultura animi*. People are morally obliged to strive to maintain simple esteem, that is, “the reputation of being a good man”. Second, the exercise of the natural law duties is advantageous for an individual’s personal well-being and vice versa. Third, people are prohibited from protecting their simple esteem violently. This implies that, rather than conceiving simple esteem as an inalienable natural right, Pufendorf is mostly concerned about the conditions of peaceful social life. This emphasis is in tune with his prevailing concern with the maintenance of political stability as a precursor of sociability. The same emphasis can be seen in the case of the natural right to one’s body and life. Violent self-defence against the sovereign is forbidden by natural law, because it “tends nearly always to be involved in grave disorders on such an occasion” (*JNG* 7.8.5/*PWSP* 238). Because of people’s ultimate moral duty to cultivate sociability, the natural right of simple esteem does not justify violent actions towards the sovereign or other citizens. While simple esteem is a right that all people have naturally, in practice it is defeasible by the more fundamental duty to maintain peace.

## Conclusion

This paper has shown that Pufendorf entertains a rather complex and nuanced view concerning how the elementary moral value of persons should be theorized in moral science. First of all, his distinction between morality and nature is a crucial starting point for the understanding of the nature of simple esteem. Pufendorf consistently maintains that simple esteem as a moral entity cannot be grounded in man’s rational capacities and social nature. Moreover, since Pufendorf’s interest in the concept of esteem is predominantly political, he is mostly interested in what role varying moral estimations play in social and civil life. The maintenance of lasting social order requires artificial social distinctions among individuals who are otherwise naturally equal. The precise level of people’s simple esteem always hinges on the moral requirements of the divinely commanded law of sociability. Yet, it is people themselves who must rationally attribute the correct level of simple esteem to themselves and others in varying social contexts. The purpose of man-made social distinctions is to create order and harmony in social life as natural law demands.

Pufendorf regards the practices of simple esteem as a natural duty imposed by natural law. To retain a certain level of simple esteem is indispensable for the fulfilment of one’s natural law duties. Other people are obliged to respect our natural right to simple esteem because we need simple esteem to fulfil our duties as a man and citizen. In a civil state,

---

<sup>11</sup> Translation slightly modified.

the level of one's simple esteem is heavily reliant on the will of the sovereign. Yet, to some extent the foundations of simple esteem are autonomous from civil laws. The sovereign cannot arbitrarily obliterate a citizen's (natural) simple esteem. Despite the absolutistic tendencies of Pufendorf's political theory, simple esteem is independent from the legitimate authority of the state and depends on one's performance as a sociable citizen. In this sense, Pufendorf's treatment of the concept of simple esteem creates the conceptual possibilities for esteem-giving and esteem-seeking as social and moral forces that become a central idea in the conceptions of the self-governing spheres of civil society in the Enlightenment.

## Bibliography

### Primary Sources

- JNG Pufendorf, Samuel. 2002. *De jure naturae et gentium. Samuel Pufendorf, Gesammelte Werke*. Band 4, edited by Frank Böhling. Berlin: Akademie Verlag.
- LNNO Pufendorf, Samuel. 1934. *Of the Law of Nature and Nations*. Translated by C. H. Oldfather & W. A. Oldfather. Oxford: Clarendon Press.
- PWSP Pufendorf, Samuel. 1994. *The Political Writings of Samuel Pufendorf*. Translated by Michael Seidler and edited by Craig L. Carr. New York: Oxford University Press.

- Hobbes, Thomas. 1991. *Leviathan*. Edited by Richard Tuck. Cambridge: Cambridge University Press.
- Pufendorf, Samuel. 1991. *On the Duty of Man and Citizen*. Translated by Michael Silverthorne and edited by James Tully. Cambridge: Cambridge University Press.

### Secondary Sources

- Darwall, Stephen. 2012. "Pufendorf on Morality, Sociability, and Moral Powers." *Journal of the History of Philosophy* 50(2): 213–38.
- Debes, Remy (ed.). 2017. *Dignity: A History*. Oxford: Oxford University Press.
- Haakonssen, Knud. 1996. *Natural Law and Moral Philosophy: From Grotius to the Scottish Enlightenment*. Cambridge: Cambridge University Press.
- Haakonssen, Knud. 2010. *Natural Law and Personhood: Samuel Pufendorf on Social Explanation*. European University Institute, Max Weber Lecture No. 2010/06. [http://cadmus.eui.eu/bitstream/handle/1814/14934/MWP\\_LS\\_2010\\_06.pdf?sequence=1](http://cadmus.eui.eu/bitstream/handle/1814/14934/MWP_LS_2010_06.pdf?sequence=1)
- Haara, Heikki. 2018. *Pufendorf's Theory of Sociability: Passions, Habits and Social Order*. Cham: Springer.
- Haara, Heikki and Lahdenranta, Aino. 2018. "Smithian Sentimentalism Anticipated: Pufendorf on the Desire for Esteem and Moral Conduct." *Journal of Scottish Philosophy* 16(1): 19–37.

- Haara, Heikki and Stuart-Buttle, Tim. 2019. "Beyond Justice: Pufendorf and Locke on the Desire for Esteem." *Political Theory* 47(5): 699–723.
- Honneth, Axel and Fraser, Nancy. 2003. *Redistribution or Recognition? A Political-Philosophical Exchange*. London and New York: Verso.
- Hruschka, Joachim. 2000. "Existimatio: Unbescholtenheit und Achtung vor dem Nebenmenschen bei Kant und in der Kant vorangehende Naturrechtslehre." *Jahrbuch für Recht und Ethik/Annual Review of Ethics* 8: 181–95.
- McBride, Cillian. 2013. *Recognition*. Cambridge: Polity Press.
- Neuhouser, Frederick. 2013. "Rousseau's Critique of Economic Inequality." *Philosophy & Public Affairs* 41(3): 193–225.
- Rosen, Michael. 2012. *Dignity*. Cambridge Mass. and London: Harvard University Press.
- Saastamoinen, Kari. 2006. "Liberty and Natural Rights in Pufendorf's Natural Law Theory." In *Transformations in Medieval and Early-Modern Rights Discourse*, edited by Virpi Mäkinen and Petter Korkman, 225–56. Dordrecht: Springer.
- Saastamoinen, Kari. 2010. "Pufendorf on Natural Equality, Human Dignity and Self-Esteem." *Journal of the History of Ideas* 71(1): 39–62.
- Tuck, Richard. 1979. *Natural Law Theories: Their Origin and Development*. Cambridge: Cambridge University Press.
- Waldron, Jeremy. 2015. *Dignity, Ranks and Rights*. Oxford: Oxford University Press.
- Zurbuchen, Simone. 2019. "Dignity and Equality in Pufendorf's Natural Law Theory." In *Philosophy, Rights and Natural Law: Essays in Honour of Knud Haakonssen*, edited by Ian Hunter and Richard Whatmore, 147–68. Edinburgh: Edinburgh University Press.



# Shame, Common Wealth and Religion in the Thought of Miklós Bethlen (1642–1714)<sup>1</sup>

## Introduction

The call for papers of the Third Budapest Seminar in Early Modern Philosophy drew the attention of potential participants to “resources from different traditions, both ‘central’ and ‘peripheral’ ones”. I will discuss “peripheral sources” in relation to my ongoing research, concerning which I have rather questions and problems instead of solid and established answers. Miklós Bethlen, the hero of this paper, wrote his texts mirroring his philosophical thoughts on the periphery of institutional conditions of the Early Modern history of European philosophy. His name is well-known for scholars of early modern Hungarian *history* (R. Várkonyi 1984, 1999) or early modern Hungarian *literature* (Ötvös 2013; Tóth 2012; Förköli 2010; Maráczsi 2005; Gömöri 2001; Czigány 1984; Tolnai 1970).<sup>2</sup> One can say without exaggerating that the whole history of Hungary and Transylvania between 1660 and 1704 can be described in the form of footnotes to his biography. His main political achievement was the so-called *Diploma Leopoldinum* (1691), which declared the legal status of the Principality of Transylvania within the Habsburg Empire after its liberation from Ottoman rule (cf. R. Várkonyi 1994). In the course of his negotiations with Kaiser Leopold, Bethlen succeeded in achieving guarantees for a much more liberal legal status than had previously been expected. The regulations of the *Diploma* proved to be enduring, they remained in place until the revolutionary events of 1848.

---

1 This research was supported by project nr. EFOP-3.6.2-16-2017-00007, titled Aspects on the development of intelligent, sustainable and inclusive society: social, technological, innovation networks in employment and digital economy. The project has been supported by the European Union, co-financed by the European Social Fund and the budget of Hungary.

2 These references include chiefly contributions in English, German or French. For a more detailed guide regarding the research, see Tóth 2016!



But the protagonist of my paper is Miklós Bethlen, the *Philosopher*<sup>3</sup> (Lacházi 2007, 2015, 2016; Simon 2016a, 2016b, 2017), who as such remained in the shadow of his role played in the local and international scenes of politics. In the present paper I would like to introduce some of the philosophical ideas developed in the *Foreword* of his own *Autobiography* (*Élete leírása magától*) written in Hungarian in 1708.<sup>4</sup> It is more than puzzling that the only text which provides us with the possibility to reconstruct his philosophical ideas was written in the very late period of his life; Bethlen remained in silence regarding philosophical matters until 1708, when after he had been arrested in 1704 accused of high treason he was transported from Nagyszeben (today Sibiu, Romania) to Vienna. Along the way he was held in prison for two months in May and June 1708 in Eszék (today Osijek, Croatia). According to his own account, he composed the *Foreword* to his *Autobiography* there, while the rest of the manuscript containing the extended narration of his own life was produced by him later in his prison in Vienna during the years from 1708 to 1710. The available sources of his own life present him as an intellectual of high quality regarding practical matters, but we have no sign of the refined philosophical orientation in the documents preceding the *Foreword* of the *Autobiography*.

Bethlen was born in 1642 into one of the most ancient families of Transylvanian high nobility. Members of his family played an important role throughout the sixteenth and seventeenth centuries of early modern Transylvania struggling in the political space of power between the Ottoman Empire and the Habsburg Monarchy, respectively. Bethlen grew up in the immediate circle of Transylvanian Princes in the exclusive group of other youths with a similar high social background. In the 1650s he attended lectures at the princely court in Gyulafehérvár (today Alba Iulia, Romania) given by János Apáczai Csere, who is honoured as the first Hungarian Cartesian thinker (Hanák 1990, 29–31).

In 1661–1663 Bethlen visited the University of Heidelberg in Germany and the Universities of Utrecht and Leiden in the Low Countries. While in 1659 Bethlen studied under the Ramist-Cartesian Apáczai Csere; shortly after the death of his teacher he attended the lectures of the young Samuel Pufendorf in Heidelberg in the early months of 1661.<sup>5</sup> Bethlen mentions his German professor in the usual laconic manner of the *Autobiography*, nevertheless he does not fail to note that Pufendorf lectured on Grotius's masterwork *De iure pacis ac belli* (Bethlen 1980, 573). One cannot overestimate the striking difference in intellectual milieu between Transylvania and Heidelberg: while attending Pufendorf's lectures, Bethlen must have gained immediate experience of his professor's most intensive dilemmas concerning *a priori*-mathematical *versus* historical argumentation for the theory

---

3 There is some research concerning Bethlen's philosophy, however neither Lendvai's (2001) nor Hanák's (1990) laconic interpretations provide us with insights into the philosophical substance of the *Foreword* in Bethlen's *Autobiography*.

4 Despite all my caveats (cf. Simon 2016b), I refer to Bethlen's text in the standard edition (1980).

5 See Benrath's publication of the advertisements of lectures at the University of Heidelberg in 1661 (Benrath 1961, 95).

of positive and natural rights, or concerning the removal of moral theology from the realm of natural rights (Dreitzel 2001, 779).

The other scene of his higher education was the Low Countries. Bethlen studied at the universities of Utrecht and Leiden between 1661 and 1663. He attended both public and private lectures of leading intellectuals belonging either to the philosophical movement of Cartesianism (Johannes de Raey, Henricus Regius – if Henry de Roy may be treated as Cartesian), or to the theological camp of Cocceianism (Frans Burman, Abraham Heidanus and Johannes Cocceius himself). Bethlen acquired his basic cultural identity as a Reformed Cartesian in Utrecht and Leiden, due to the decisive role played by intellectuals in the Low Countries in Cartesian philosophy joining the main stream of contemporary philosophical reflections.

Now let us turn to the *Foreword* of his *Autobiography*, written more than 40 years after having finished his university studies! I divide my paper into four parts. First, I present Bethlen's mechanical interpretation of speech acts; it follows the discussion of the influence of Pufendorf's teaching about moral entities on Bethlen's conception of common wealth. Thirdly, I will touch on the function of shame within and outside the positive state of political community and introduce Bethlen's ideas concerning the religious inclinations of humans as the basis of natural law. Last, but not least, I will highlight some questions and problems which emerge in the constellation of these seemingly divergent topics in Bethlen's text.

I will not discuss philological problems. I will quote Bethlen's text according to my own research to reconstruct an *ultima manus* version of his *Foreword*, which in some cases proves to be different from the formulations of the standard modern edition (cf. Simon 2016b). For the English version I follow Bernard Adam's translation published in London in 2004,<sup>6</sup> but I made emendations of some passages relevant to our topic.

## Mechanism of Speech Acts

According to Bethlen, we are living in a physical world which is mechanical in its entire nature. He accepts many of Descartes' teachings, such as the threefold division of corpuscular particles as developed in *The Principles of Philosophy*.<sup>7</sup> There are minds merged into mechanical bodies, or at least we anticipate the existence of immaterial souls while perceiving social behaviour and speech acts of others. However, our anticipation turns out to be false most of the time. Bethlen is deeply convinced that social behaviour and communicative speech acts can be generally described in a mechanical way. This means that ordinary cases of behaviour in everyday life lack consciousness. One is subjected to various physical effects in a society, to which he or she gives automatic responses without being conscious

---

6 *The Autobiography of Miklós Bethlen*, transl. by Bernard Adams, London, Kegan Paul, 2004.

7 Descartes AT VIII/1, 80–202, passim, esp. 103–15.

of it. This also holds in the special case of communicative speech acts. In sharp contradiction to Descartes' intention,<sup>8</sup> Bethlen claims that most cases of speech performances in a society prove to be mechanical and are not indubitable signs of the existence and operation of immaterial mind (Bethlen 1980, 434–6).

Does not even the simplest of men experience this within himself; if his mind is fixed on something, he knows nothing about whatever he may hear or see apart from that which he then perceives. Indeed, he often does not know whether or not he has said that to which his tongue is accustomed; he may say a prayer or whatever, but his mind is far away. Sometimes he will even speak to another and answer, but he will not know of it when questioned later, and it will merely seem like a familiar dream that comes to mind. This everyone may experience best of all in church, during divine service, when repeating a prayer after the priest or listening to a sermon; if one's soul is far away one will know nothing of what one's tongue is saying or one's ears are hearing, but one will be perfectly well aware of that which one's mind has perceived. (Bethlen 1980, 434; 2004, 52)

Bethlen's example of the automatically praying man whose mind is intended to an entirely other object than the expressed content and whose mind directs the bodily organs for performing speech act without consciousness probably follows Claude Cleselier's *Preface* to the Paris edition of Descartes' *De l'homme* in 1664. There are several problems concerning this account of unconscious use of language. Without going into details,<sup>9</sup> for the purpose of the present paper it is enough to cognise that Bethlen's treatment of the topic depends on Cartesian discussions in France and in the Netherlands of the 1660s and 1670s.

## Pufendorbian Moral Entities<sup>10</sup>

Let's take a new point of departure! Bethlen develops a theory of reputation, esteem<sup>11</sup> and honour in a community. Members of a society always evaluate themselves and others, always make judgements of their own or of other's good or bad deeds, intentions and behav-

---

8 Descartes AT VI, 57–59.

9 See my interpretation, Simon 2017.

10 Cf. Dreitzel 2001, 779–81 and Heikki Haara's contribution to the present volume!

11 The Hungarian equivalent for Pufendorbian *existimatio* is *hírnév*, a compound word from *hír* (rumour, news) and *név* (name). Consequently, Bethlen divides his definition of esteem into two parts following the structure of the Hungarian compound. Bethlen's definition of *name* as a linguistic unit focuses on its communicative character instead of treating it primarily as an expression of mental content. The mechanical structure of a word as sign is determined by its communicative effectivity. On the other hand, communicative effectivity is made responsible for creating Pufendorbian civil and moral entities by Bethlen.

our respectively. By doing so they can judge correctly or commit errors – paradigmatically we cannot even make appropriate judgments of our own behaviour, that is, we esteem ourselves erroneously. However, whether these evaluations are correct or not, reputation and esteem produce special social entities, which Bethlen calls moral or civic qualities (*qualitates morales sive civiles*; Bethlen 1980, 410). Bethlen speaks about qualities that legitimate political authority, that make the representants of different kinds of practical knowledge to be honoured in a society and that construct relationships within families. The idea is clearly Pufendorffian in its origin, especially as Bethlen emphasises the ontological efficacy of these qualities (“they can become real and acts affecting men” – Bethlen 1980, 411; Bethlen 2004, 26). It is important to note that Bethlen is aware of the Pufendorffian principle that esteem itself is not a civil or moral quality, but it *produces* civil and moral qualities. As mentioned above, Bethlen refers laconically to the young Pufendorff’s lectures on Grotius, which he attended in Heidelberg in the year 1661 (Bethlen 1980, 573). The theory concerning the ontological status of moral entities was already involved in the 1660 publication of the German scholar’s *Elements of Jurisprudence* (Böhling 2014, 10–2).<sup>12</sup> Bethlen’s use of Pufendorff’s theory is quite loose. While alluding to specific *moral qualities* imposed on things and persons by esteem and speech performances, Bethlen does not refer to the general concept of moral *entities*, of which moral *qualities* are only but one type besides moral *state*, moral *substances* and moral *quantities*. Without making it explicit, he seems to accept that there is a moral state of humans imposed on physical space, in which moral substances are imposed on physical bodies.

Now, with the connection of both the respects of mechanical interpretation of speech acts and the ontology of moral entities, we arrive at a very strange point that cannot be found in Pufendorff. Namely, Bethlen stresses that speech acts play a crucial role in the genesis and reality of moral qualities. Against this background, esteem and honour are paradigmatically results of unconscious verbal interactions in society; therefore they can be described in a mechanical way as well.<sup>13</sup>

---

12 Cf. Dreitzel’s note that the doctrine of *entia moralia* grew out of Pufendorff’s “Auseinandersetzung” with Erhard Weigel’s “specific Cartesianism” (Dreitzel 2001, 780). On the relationship between Weigel and the young Pufendorff, cf. Wolfgang Röd’s classic summary: Röd 1969.

13 In this sense my interpretation stands very close to what Haara and Lahdenranta call “sentimentalism”: “By sentimentalism, we mean a metaethical view that takes moral qualities to be constituted by human judgements which either are or have their basis in emotional responses.” (Haara-Lahdenranta 2018, 22) Bethlen’s specific contribution to the relationship between moral-civil qualities and human judgments consists in the incorporation of the Cartesian dilemma of (un)consciousness of speech acts.

# Shame Within and Outside of Community

Bethlen introduces the topic of shame as follows:

From these [i.e. from one's esteem/honour from the perspective of others – JS] there arise and born naturally in a man two things: *ambitio* or *generositas*, *gloriae cupido*, which is the desire for favourable judgement upon him; and shame, disgrace, *pudor*, which is fear and horror of unfavourable judgement. (Bethlen 1980, 413; 2004, 28)

Shame and ambition<sup>14</sup> are emotions which reflect on intersubjective esteem and honour among people. Shame is a reflective prolongation of mechanical structure of social qualities on the level of emotions, its aim is a kind of emotional management of social esteem; that is, a strategy to avoid actions which result in negative evaluations of others in my direction. However, shame as one's psychological disposition does not depend on whether the evaluative judgment is correct or false. May my fellows in a society make a *false* judgement of my intended act, I will intend *not* to perform the act in question, that is I will feel shame to do so *as if* the expected evaluations of my fellows in a society were *right*. The certainty of my conscience's eventually correct judgement about myself is not so strong as it were able to eliminate my natural felt shame caused by the expected unfavourable judgement of others, even if this latter is false. The same with ambition: Even if the morsel of conscience informs me correctly concerning the evilness of my planned act in the future, I will intend to perform it ambitiously, because of the expected favourable judgment of others in society.

The efficacy of shame and ambition in a community presupposes a general view of the sociality of human beings. Despite all the similarities of Bethlen's theory concerning reputation, esteem and honour with Pufendorf's teaching, the Hungarian thinker builds his interpretation of society on motifs other than those of his professor in Heidelberg. Bethlen explains the basic feature of the sociality of humans in terms of their shared anthropological standard of religious inclinations. Paraphrasing the Ciceronian-Senecan *consensus* argument in the manner of Gerhard Vossius (Vossius 1668, 1–2) and Herbert of Cherbury (Herbert of Cherbury 1663, 2), Bethlen establishes a general leaning towards religious behaviour in all nations (Bethlen 1980, 427–432; 2004, 45–50). He introduces three moments as anthropological standards and builds an idea of natural religion on it: 1) natural inclination towards esteem and honour even after one's death, 2) belief in God and 3) conscience. God implanted these "*qualities or inclinations*" in man through nature – the anthropological weight of these natural features is so strong that even original sin could not extinguish or modify them. They are also present in all pagan nations, who "prefer to worship stones, trees, reptiles or frogs than to be without a God to whom to turn in time of need" (Bethlen

---

<sup>14</sup> For the prevalence of shame and ambition as emotional constituencies of Pufendorfsian human sociality, cf. Haara 2018, 99–136!

1980, 427; 2004, 45). However, Bethlen did not follow the typical procedure of comparative studies of religions of his time: instead of describing different religions in the light of their worships of physical objects, the argumentation focuses on intersubjective relations resulting from speech performances in a *social* context. It follows consequently from the basic religious feature of mankind that none of the nations lacks and lacked social phenomena depending on speech acts: “[t]here is no nation on earth in which there is not worship, oath-taking, blessing, cursing [átok], reproach [szitok], some slight modesty, glory, outrage, reward and punishment” (Bethlen 1980, 427; 2004, 45). Therefore, Bethlen embeds Pufendorfian civil and moral constituents of society into a theory of natural religion.

But one may suppose that these latter phenomena exist only under the circumstances of a positive state of commonwealth: “Now let us suppose that glory, outrage, reward and punishment are entirely human inventions for the maintenance of society.” (Bethlen 1980, 427; 2004, 45) Bethlen expresses the volatile and unstable feature of these social emotions through describing *reproach* as “the blazing up the angry heart, as it were the burning of a sooty chimney” (Bethlen 1980, 427). Here, we have the idea discussed above: inclination to glory is ambition towards expecting reward and esteem because of acting in accordance with the interest of the common good, or as Bethlen has it “maintaining of society”. *Outrage*, on the other hand, is punishment for violation of the interest of the common good. May these emotions have their origin deep in human nature, the content of the common good directing them in a social context proves to be contingent in its very nature, or as Bethlen writes: they “are entirely human inventions for the maintenance of society”. But mere fear of punishment in the case of violating the common good cannot account perfectly for feeling shame; and hope of reward in the case of acting in accordance with the common good cannot account perfectly for keeping promises in oath-taking.

Would there be worship, oath-taking, blessing and cursing if man did not believe in God, who has the power and the will, if you pray, to listen and come to your aid; to punish him that swears falsely; to implement blessing and cursing? For surely everyman knows that that is all merely words, the motion of the tongue and the air, and if there is none to give it power it is all futility and the wind bears it away. Would he that swears have any regard for that, or he to whom he swears, he that blesses or curses or those to whom these are directed? would they make use of them? would they fear them? if they had no faith of them. Would the powerful man who fears no punishment blush and feel ashamed where no one can see him, if his conscience did not prick him? (Bethlen 1980, 427–428; 2004, 45)

While leaving the positive social context behind, Bethlen arrives at the domain of natural laws. Bethlen appeals to God as the condition of causal efficacy of speech acts like oath-taking, blessing or cursing. However, it does not mean that the interpretation of their normativity beyond the positive legal state relies on supernatural moments of belief. Inclination towards pursuing glory or inclination towards avoiding shameful actions are natural endowments of mankind. Addressing the Hobbesian figure of the *powerful* in the

natural state, Bethlen declares that he feels shame even if knowledge of society, negative evaluation and the prospect of punishment for violating the common good are suspended. Therefore, shame proves to be not only an emotional management of esteem and honour in a community, but an emotion which mirrors a non-conventional, natural order of values. In this sense, shame has a strong connection to conscience. Even if Bethlen introduced conscience as a natural faculty of humans with theistic origin without the condition of revealed religion, one might ask whether feeling shame beyond and independently of the conventional positive state of a community has its origin in the anthropological standard of original sin.

Bethlen, being a convinced Calvinist, is aware of this challenge of Augustinian theological anthropology, and addresses the relationship between shame and original sin directly:

From that it is immediately apparent that the words of Moses are no *fabula*, and clothing is Adam's invention not only against warm or cool but also against shame. [...] In vain and falsely have certain atheists and profaners postulated that Moses and others like him wished to obtain honour for themselves, domination over others and the obedience of the simple, and therefore invented God, the soul, heaven, hell etc. (Bethlen 1980, 428; 2004, 45–46)

One should not be misguided by Bethlen's apology of Mosaic narration against allegedly atheistic or profane interpretations.<sup>15</sup> His vision of the relationship between shame and original sin is entirely rational and that of Enlightenment. It is not the Mosaic law which provides us with a theological interpretation of natural shame, but it is the rational theory of naturally felt shame which justifies the *truth* of the inspired Mosaic narration of Adam's fall. In other words: we do not feel shame beyond social conventions because we are heirs of the theologically established original sin, but we as Christians inherit original sin because we, similar to representants of other confessions, are naturally ashamed to perform some acts. Bethlen explores a thin area for shame as natural emotion between the robust interpretative perspectives of social psychology and theological anthropology. He does not deny that one may feel shame because he or she experiences dishonour from his or her fellow citizens, and he does not rule out the possibility that one has bad conscience while feeling shame. But it is rather the natural emotion of shame, which allows for shame as an individual strategy for avoiding dishonour in society, as well as for upholding the *truth* of the Mosaic narrative of Adam's feeling shame after the Fall.

At this point, we can contrast Bethlen's formulations with Pufendorf's ideas again. The German scholar addresses the problem of deduction of natural laws' normativity from

---

15 For a defence of natural law theory against atheism, cf. DJN, 3,4,4 (Pufendorf 1672, 312–4). Pufendorf's main target is Hobbes's discussion of the atheist position, according to which there is no system of natural values that the atheist could violate. Therefore, Hobbes maintains that atheists do not violate natural law, but commit the error of ignorance and must be punished according to the jurisdiction of war instead of natural law.



theologically established shame in *Elements of Jurisprudence*, Book 1, definition 12, paragraph 15 directly.

Now there is controversy among the learned about the strength of obligation in pacts which have been formed by the law of nature alone, such as exist among those who acknowledge no common judge in a human forum, or about which the civil law makes no disposition. For some have maintained that the efficacy of such pacts consists in a bond of shame and modesty [*pudoris ac verecundiae vinculo*] alone, especially where no agreement [*συνάλλαγμα*] has as yet interceded and nothing has been furnished by either side, while the rest harshly criticize this opinion as weakening the trustworthiness of all treaties.

To us it seems that the matter is not so difficult if it is initially supposed that men have been fashioned by nature to cultivate society among themselves, and that no one should inflict on another that which can furnish a cause for discords and wars [...] it is quite apparent that men are altogether obligated by the law of nature to observe their pacts and that those who violate them sin against it [...] .<sup>16</sup> (Pufendorf 1994, 55)

Pufendorf's idea of sociability as an anthropological standard of mankind makes the normative function of shame for acting in accordance with the natural law unnecessary. Bethlen does not admit human sociability even in its Pufendorffian version of ability to impose moral entities upon physical states of affairs and to be aware of natural values originating from these impositions. He insists on the normative function of shame for respecting natural rights of others and disqualifies human rationality for recognising natural values. Bethlen seems to occupy a third position within Pufendorffian-Hobbesian ramification of social norms. He accepts the Pufendorffian theory of civil entities and ascribes causal efficacy to them in a society; but this latter ability of introducing physical and mental changes in the world is bound to the mediation of psychological states and speech acts in human beings. Bethlen's sensitivity to psychological aspects of social norms and limiting the consciousness of speech acts converges in a Hobbesian direction, although the English thinker would have accepted neither Bethlen's vision of universal normativity of naturally felt shame in the natural state of humans, nor the religious syncretism built on it.

---

<sup>16</sup> Latin and Greek insertions from Pufendorf 1660, 106–7. Cf. Pufendorf 1686, chapter 5, paragraph 5, 276–7.



## Summary and Problems

There are some puzzling moments and problems in Bethlen's discussion of the relationship between shame, commonwealth and religion. If we accept Bethlen's view of shame as a reflective psychological management concerning acquiring esteem in a positive state of community, the religious aspect of shame proves to be problematic. According to the Hungarian author, natural religion (as a common and latent background behind all confessions) is based on a natural inclination towards reputation, esteem and honour – besides a reductive knowledge of God and the presence of conscience. The link between social and natural shame remains unexplained in the *Foreword* of the *Autobiography*. For maintaining natural shame outside human community, Bethlen should extend the relevance of Pufendorfian moral qualities to the realm of the natural state of humans as well. Bethlen seems to substitute natural – or as Pufendorf describes it, *simple* – esteem with the basic drive for religious behaviour in the human natural state. Furthermore, as Bethlen makes esteem strongly dependent on speech acts in the social context, he should treat the problem of use of language in the natural state, too. Are communicative performances in most cases *unconscious* in the *natural* state of humans as well, or is the mind *aware* of the contents of speech in the state of nature while producing natural moral qualities? There are some signs towards this latter direction, but this step is in no way explicit in Bethlen's text. The lack of this connection is crucial for his theory, because he seems to base natural religion on a natural inclination towards esteem.

But one thing is clear. Despite all the inconsequencies, we might suppose in Bethlen's application of Pufendorf's theory of moral qualities a deep understanding of the problems concerning human community. The gap between moral qualities and natural religion as explanatory ideas for human sociality is abridged by Bethlen's introduction of the Cartesian problem of the mind's transparency concerning speech acts. In Bethlen's vision of society, Pufendorfian imposition is driven through religious inclinations of human nature on the basis of mechanical speech performances. However, Pufendorf had no interest in embedding his theory of imputation into the problems of the Cartesian philosophy of mind. Likewise, prominent protagonists of Cartesianism who subscribed to the possibility of unconscious speech acts (such as Claude Clerselier (Clerselier 1664, oo ii<sup>v</sup>), Jean Darmanson (Darmanson 1684, 62) or Jean-Baptiste Duhamel (Duhamel 1673, 89–90)), appeared to be entirely disinterested in the Pufendorfian enterprise of political philosophy. One possible way to bridge the gap between these tendencies can be found in Thomas Hobbes's philosophy. There may have existed parallel phenomena in Pufendorf's Early Modern reception of Bethlen's reformed Cartesian interpretation – above all in the Protestant Low Countries and Germany – which should be explored in the future.

# Bibliography

## Primary Sources

- Bethlen, Miklós. 1980. “Élete leírása magától.” In *Kemény János és Bethlen Miklós művei*, edited by V. Windisch Éva, 399–981. Budapest: Szépirodalmi.
- Bethlen, Miklós. 2004. *The Autobiography of Miklós Bethlen*. Translated by Bernard Adams. London: Kegan Paul.
- Clerselier, Claude. 1664. “Preface.” In René Descartes. *L’Homme de René Descartes et un traité de la formation du foetus du mesme autheur. Avec des Remarques de Louys de La Forge*, e–oiiii. Paris, Charles Angot.
- Descartes, René. 1664. *L’Homme de René Descartes et un traité de la formation du foetus du mesme autheur. Avec des Remarques de Louys de La Forge*. Paris, Charles Angot.
- Descartes, René. 1902: *Discours de la méthode*. Oeuvres de Descartes. Publ. par Charles Adam et Paul Tannery, vol. 6. Paris: Cerf.
- Descartes, René. 1905: *Principia Philosophiae*. Oeuvres de Descartes. Publ. par Charles Adam et Paul Tannery, vol. 8. Paris: Cerf.
- Herbert of Cherbury. 1663. *De religione gentilium, errorumque apud eos causis*. Amstelodami: Blaeu.
- Pufendorf, Samuel. 1660. *Elementorum iurisprudentiae universalis libri duo*. Hagae–Vomitis: Vlacq.
- Pufendorf, Samuel. 1672. *De jure naturae et gentium libri octo*. Lundt: Junghans.
- Pufendorf, Samuel. 1686. *Eris scandica qua adversus libros de Iure Naturali et Gentium obiecta diluuntur*. Frankfurt am Main: Knochius.
- Pufendorf, Samuel. 1995. “Elements of Jurisprudence.” In *The Political Writings of Samuel Pufendorf*, edited by Craig L. Carr, translated by Michael J. Seidler, 29–92. New York – Oxford: Oxford University Press.
- Vossius, Gerardus. 1668. *De theologia gentili et physiologia christiana sive de origine ac progressu idololatriae [...] Libri IX*. Amstelodami: Blaeu.

## Secondary Sources

- Benrath, Adolf. 1961. “Heidelberger Vorlesungsverzeichnisse aus den Jahren 1655, 1658 bis 1662 und 1685.” *Heidelberger Jahrbücher* 5: 85–103.
- Böhling, Frank. 2014. “Einleitung.” In Samuel Pufendorf: *De jure naturae et gentium. Dritter Teil: Materialien und Kommentar von Frank Böhling*. Gesammelte Werke. Hrsg. von Wilhelm Schmidt-Biggemann, Band 4/3. Göttingen: De Gruyter.
- Czigány, Lóránt (ed.). 1984. *The Oxford History of Hungarian Literature: From the Earliest Times to the Present*. Oxford, Clarendon Press.
- Darmanson, Jean. 1684. *La beste transformée en machine*. [s.l.]
- Dreizel, Horst. 2001. “Samuel Pufendorf.” In *Grundriss der Geschichte der Philosophie. Band 4, Das Heilige Römische Reich deutscher Nation – Nord- und Ostmitteleuropa*.

- Hrsg. von Helmuth Holzhey und Wilhelm Schmidt-Biggemann, Band 4/2, 757–812. Basel: Schwabe & Co AG Verlag.
- Duhamel, Jean-Baptiste. 1673. *De corpore animato libri quatuor*, Paris, Michallet.
- Förköli, Gábor. 2010. “A taláros francia és az erdélyi főúr: Jacques-Auguste de Thou *Commentarii de vita suája* mint Bethlen Miklós *Önéletírásának* előképe” [The French in Robe and the Transylvanian Baron: Jacques-Auguste de Thou’s *Commentarii de vita sua* as a model of Miklós Bethlen’s *Autobiography*]. In *Epika: Fiatal kutatók konferenciája*, edited by Dobozy Nóra Emőke et al., 23–36. Budapest: ELTE BTK Régi Magyar Irodalom Tanszék.
- Gömöri, George. 2001. “Miklós Bethlen.” In *Encyclopedia of Life Writing: Autobiographical and Biographical Forms*, edited by Margaretta Jolly, 102–3. London-Chicago: Fitzroy Dearborn Publishers.
- Haara, Heikki. 2018. *Pufendorf’s Theory of Sociability: Passions, Habits and Social order*. Cham: Springer.
- Haara, Heikki and Aino, Lahdenranta. 2018. “Smithian Sentimentalism Anticipated: Pufendorf on the Desire for Esteem and Moral Conduct.” *Journal of Scottish Philosophy* 16(1): 19–37.
- Hanák, Tibor. 1990. *Geschichte der Philosophie in Ungarn. Ein Grundriß*. München: Trofenik.
- Lacházi, Gyula. 2007. “Szenvedély, jellem és morál Bethlen Miklós önéletírásában” [Passion, Character and Moral in Miklós Bethlen’s *Autobiography*]. In *Emlékezet és devóció a régi magyar irodalomban*, edited by Balázs, Mihály and Gábor, Csilla, 247–58. Kolozsvár: Egyetemi Műhely Kiadó.
- Lacházi, Gyula. 2015. “A semmi Bethlen Miklós antropológiájában és önéletírásában” [The Notion of Nothing in Miklós Bethlen’s Anthropology and Autobiography]. *Irodalomtörténeti Közlemények* 119(3): 336–55.
- Lacházi, Gyula. 2016. “A magyarországi korai felvilágosodás lélektani irodalma és Bethlen Miklós önéletírásának lélektani relevanciájú megállapításai” [The Psychological Literature of Hungarian Early Enlightenment and Psychological Statements in Miklós Bethlen’s *Autobiography*]. In *A felvilágosodás előzményei Erdélyben és Magyarországon (1650–1750)*, edited by Balázs, Mihály and Bartók, István, 217–26. Szeged: SZTE Magyar Irodalmi Tanszék.
- Lendvai, Ferenc F. 2001. “Siebenbürgen.” In *Grundriss der Geschichte der Philosophie. Band 4, Das Heilige Römische Reich deutscher Nation – Nord- und Ostmitteleuropa*. Hrsg. von Helmuth Holzhey und Wilhelm Schmidt-Biggemann, Band 4/2, 1383–8. Basel: Schwabe & Co AG Verlag.
- Marácz, Géza. 2005. “Miklós Bethlen in 17<sup>th</sup> Century England: with the English Translation of His Travel Account.” *Irodalomismeret* 16(1): 98–102.
- Ótvös, Péter. 2013. “Memoirenliteratur: Miklós Bethlen (1642–1716).” In *Geschichte der ungarischen Literatur. Eine historische-poetologische Darstellung*, hrsg. von Ernő Kulcsár Szabó, 83–6. Berlin-Boston: De Gruyter.

- R. Várkonyi, Ágnes. 1984. *Erdélyi változások. Az erdélyi fejedelemség a török kiűzésének korában 1660–1711* [Transylvanian Changes: The Principality Transylvania in the Period of the Expulsion of the Turks 1660–1711]. Budapest: Magvető.
- R. Várkonyi, Ágnes. 1994. “The End of Turkish Rule in Transylvania and the Reunification of Hungary (1660–1711).” In *History of Transylvania*, edited by Köpeczi, Béla et al., 359–411. Budapest: Akadémiai Kiadó.
- R. Várkonyi, Ágnes. 1999. “Bethlen Miklós és II. Rákóczi Ferenc angol–holland kapcsolataihoz [English and Dutch Political Connections of Miklós Bethlen and Ferenc Rákóczi II.]. *A Ráday Gyűjtemény Évkönyve* 9: 3–17.
- Röd, Wolfgang. 1969. “Erhard Weigels Lehre von den entia moralia.” *Archiv für Geschichte der Philosophie* 51(1): 58–84.
- Simon, József. 2016a. “Bethlen és a hírnév. A Pufendorf-recepció problémája” [Bethlen and Esteem. The Problem of Pufendorf’s Reception]. In *A felvilágosodás előzményei Erdélyben és Magyarországon (1650–1750)*, edited by Balázs, Mihály and Bartók, István. Szeged: SZTE Magyar Irodalmi Tanszék, 187–205.
- Simon, József. 2016b. “Filológiai és filozófiatörténeti megjegyzések Bethlen Miklós Önéletírásának Előljáró beszédéhez” [Philological and Philosophical Comments on the Foreword of Miklós Bethlen’s *Autobiography*]. *Irodalomtörténeti Közlemények* 120(3): 299–314.
- Simon, József. 2017. “Bethlen Miklós és az imádkozó automata” [Miklós Bethlen and the Praying Automaton]. *Magyar Filozófiai Szemle* 61(4): 147–67.
- Tolnai, Gábor. 1970. “Miklós Bethlen un classique des anciens mémoires hongrois.” *Acta Litteraria Academiae Scientiarum Hungaricae* 12(3–4): 251–72.
- Tóth, Zsombor. 2016. *Bethlen Miklós – válogatott bibliográfia* [Miklós Bethlen – A Selected Bibliography]. Budapest: reciti 2016.
- Tóth, Zsombor. 2012. “A Man for All Seasons: Exile, Suffering and Martyrdom in the Autobiography of Miklós Bethlen.” *Hungarian Studies* 26(2): 273–83.

## Love and Order: Malebranche and the Feeling of Natural Law

In *An Enquiry Concerning the Principles of Morals* David Hume discusses the notion of justice and its grounds and favourably cites Montesquieu for recognizing that laws should depend on each society's particular situation. Nonetheless, Hume's appreciation of the author of *The Spirit of the Laws* stops here. As he remarks in a footnote,

This illustrious writer (= Montesquieu), however, sets out with a different theory, and supposes all right to be founded on certain *rappports* or relations; which is a system, that, in my opinion, never will be reconciled with true philosophy. (Hume 1983, 29)

Far from being an ally against the school of natural law, Montesquieu's doctrine of law, in Hume's opinion, is imbued with rationalism, which is a legacy of the French philosopher: Nicolas Malebranche – “the first that started this abstract theory of morals” which “excludes all sentiment, and pretends to found everything in reason” (Hume 1983, 29). However, is this true? Does Malebranche's ethics represent a rationalistic model of understanding human moral behaviour or is it possible to read it differently? In the following paragraphs, I will dwell upon this question to show through an analysis of Malebranche's reflection on the role played by humans' affective dimension in morality, that Hume's judgement is not completely fair. From this inquiry, it will also emerge that Malebranche's arguments can be useful even in understanding the shift from the rationalist to the sentimental approach to morality in the eighteenth century. As we will see, Malebranche's analysis of human nature, by stressing more and more the role of pleasure in moral behaviour, attains a *naturalistic* description of agency, which the moral sense theory, especially Hume, will develop and deepen.

# Order, Love and Pleasure

As has been previously noted (cf. Riley 2003, 79–80), Hume’s link between “this abstract theory of morals” and Montesquieu’s statement that “all right [has] to be founded on certain *rappports* or relations” allows one to suppose that Hume has Malebranche’s definition of “Order” in mind, the “eternal law” that even God must follow. The first definition of Order appears in 1678, in the *Tenth Elucidation* of the *Recherche de la verité*,<sup>1</sup> in which Malebranche clarifies the grounds of his theory of knowledge, the vision in God. Recalling an Augustinian argument,<sup>2</sup> Malebranche states that since the truths that we (can) grasp are universal, immutable, eternal and infinite, they cannot belong to particular beings such as ourselves, and be the object, or content, of our particular reason. As a consequence, “Reason in which all men participate [...] is certainly not different from God’s own reason, for only the infinite and universal being contains in itself an infinite and universal reason” (SAT 614). Although God’s reason is “consubstantial” with God, it is in some way independent of him, since “He has to consult and follow it” (SAT 614). In this sense,

[w]e see clearly that God cannot punish innocence, that He cannot subject minds to bodies, that He is constrained to observe order. We see, then, the rule, the order, the reason of God – for what wisdom other than God’s could we see when we dare to say that God is constrained to follow it? (SAT 614)

Order is identified here with the reason of God, which contains all immutable and eternal truths.<sup>3</sup> Whereas the theoretical truths are but “relations of equality or inequality” between the perfections of God (i.e. God’s ideas),<sup>4</sup> practical truths grounded in Order

---

1 Hereinafter RDV. The reference edition for Malebranche’s works is Malebranche 1958–1967, which I will refer to as OC followed by the tome number and the page number. The *Recherche* is in OC I–III. The translation quoted in the text is Malebranche 1997. In the text the references will be SAT (= Search after Truth) followed by the page number.

2 “If we both see that what you say is true, and if we both say that what I say is true, where is it, I ask you, that we see this? Certainly, I do not see it in you, and you do not see it in me, but both of us see it in the unchangeable truth itself, which is above our minds.” Augustine, *Confessions*, Book 12, chapter 25, quoted by Malebranche in SAT 613.

3 As André Robinet has shown (cf. Robinet 1979), Malebranche capitalizes the word “Order”, especially in the first edition of the *Elucidations* to RDV (1678), to underline its divine nature. In other works, like the *Treatise on Ethics* (1684), we can find both occurrences. I use the capitalized word in the main text to highlight its technical (and Malebranche’s) use.

4 Malebranche defines God’s ideas of the creatures as “the perfections of God that correspond to these creatures, and that represent them”, SAT 68. This definition has been present since the first edition of *Recherche de la verité*. In the following editions and in his mature writings Malebranche will speak of God’s ideas in the same terms as Aquinas, that is, as the same essence of God as can be participated in by creatures. See *Summa theologiae*, q. 15, art. 2, ad 2 resp. Since, according to Malebranche (and Aquinas), God’s ideas are identical with the essence of God, it is very difficult to accept Robinet’s claim according to which Malebranche would have shared Descartes’ doctrine of the creation of eternal

derive from the difference in *degrees* of perfection, which is a qualitative difference between God's ideas:

the perfections in God that represent created or possible beings are not all equal insofar as they represent these beings, and that those, for example, that represent bodies are not as noble as those that represent minds, and furthermore, that even among those that represent only bodies or only minds, there are infinite degrees of perfection [...]. If it is true, then, that God, who is the universal Being, contains all beings within Himself in an intelligible fashion, and that all these intelligible beings that have a necessary existence in God are not in every sense equally perfect, it is clear that there will be a necessary and immutable order among them, and that just as there are necessary and eternal truths because there are relations of magnitude among intelligible beings, there must also be a necessary and immutable order because of the relations of perfection among these same beings. (SAT 618)

Malebranche explains that Order is not a mere "speculative truth" (SAT 618), which we can contemplate in a state of indifference. On the contrary, Order has "the force of law" (SAT 619) because, just as laws affect our behaviour, Order impacts our will. According to Malebranche, the normative power of Order does not derive immediately from our access to the reason of God, or contemplating the relationships between his perfections, but from God's attitude towards Order: love. Since God "loves Himself with a necessary love" (SAT 619), which is a love that inevitably flows from the full knowledge of the degree of perfection of a being, God loves beings represented through his perfections according to their degree of perfection. As a consequence,

if we wish to suppose an intelligible mind to be a thousand times more perfect than an intelligible body, the love by which God loves Himself would necessarily be a thousand times greater for the intelligible mind than for the intelligible body. (SAT 619)

In this sense, God is not free to love a stone more than a goat, because this would imply going against his same essence, in which the Order (i.e. the hierarchy of beings) is grounded. This means that in God there is no space, so-to-speak, between the *sollen* and the *müssen*, and every divine action is morally perfect because the will of God is always determined by Order.

According to Malebranche, for human beings things are a little different. Although Order also has the "force of law" for humans, "[f]or, since God has created us in His image and likeness, He cannot will that we love more what deserves to be loved less" (SAT 619)

---

truths and stated in 1674 that God created human ideas. See Robinet 1965, 207–32. On the relationship between Malebranche's and Aquinas's definition of ideas, see Priarolo 2017.



– contrary to God’s will, Order does not move our will necessarily. The reason seems to be that human beings possess two different forms of love: “natural love” and “self-love” (SAT 619). Whereas the first love drives us to follow Order, because it directs us toward God, the latter – which, in the *Tenth Elucidation*, Malebranche identifies with the Augustinian *concupiscence* – leads us against it, even though it cannot completely overcome natural love, even in damned people:

in spite of concupiscence, which conceals Order from us and prevents us from following it, Order is always a law that is essential and without exception with regard to us, and not only with regard to us but to all created intelligences and even the damned – for I do not think that they are so removed from God that they do not yet have some faint idea of Order in which they find some beauty and that they are not perhaps even ready to conform themselves to it in certain particular cases where their self-love is not at stake. (SAT 619)

To understand this passage, it should be noted that Malebranche conceives “natural love” as the movement of one’s will towards God, which, according to the doctrine of occasionalism, is the cause of every movement, and then every action, of created beings.<sup>5</sup> In this sense, “natural love” depends on the “inclination for the good in general” (RDV, IV, 1, IV, SAT 268) that constitutes our will, so that we can say “this love is but our will” (RDV, IV, 1, III, SAT 267).<sup>6</sup> As Malebranche states in the fourth book of the *Recherche*:

there is properly but one love in God, which is His love for himself, and as God can love nothing except through this love (since God can love nothing except in relation to Himself), so God imprints but one love in us, which is the love of the good in general, and we can love nothing except through this love (since we can love nothing that is not, or that does not seem to be, a good). (RDV, IV, 1, III, SAT 266–7)

After having stated that the inclination towards “the good in general” is the condition of every movement of the will, Malebranche explains that in addition to this first general inclination, we have two other main inclinations: 1) the “inclination for the preservation of our own being”, i.e. *self-love* (*amour propre*), and 2) the “inclination toward other creatures” (RDV, IV, 1, III, SAT 268).<sup>7</sup> Malebranche further divides self-love into “love of being and of well-being” (RDV, IV, 5, II, SAT 288). Whereas the first regards all desires to be superior to

---

5 Regarding occasionalism, the literature is certainly huge. See, for instance, Nadler 2010; Nadler 1993; Perler and Rudolph 2000; Lee 2008; and Favaretti Camposampiero, Priarolo and Scribano 2018.

6 In this sense we can say that love and will are coextensive and therefore one is the reciprocal of the other. See on this Vieillard Baron 1996.

7 For a detailed analysis of Malebranche’s doctrine of inclinations, see Moriarty 2006, 249–71. I would like to thank Heikki Haara for having drawn my attention to Moriarty’s works.



others and independent from God, the latter directly concerns what motivates our actions. According to Malebranche, every human being is moved by pleasure, “since pleasure is the mode [*maniere d’être*] of being that is the best and most agreeable to the soul” (ibid.). As Malebranche writes in a later edition of the *Recherche*, after a discussion with the Benedictine François Lamy on the doctrine of “pure love”,<sup>8</sup> pleasure is “the motive” of love:

it appears certain to me that pleasure in general, in so far as it contains both rational and sensible pleasures, is the unique motive or principle of natural love or of all the soul’s impulses toward any possible good, for one can love only what is pleasing. (RDV, IV, 5, II, SAT 288)

According to Malebranche, it is impossible to love something (or someone) and not to feel pleasure in it (or him/her). A sensible experience is then concerned in every act of the will, both good acts, when pleasure results from a movement toward a real good, and bad acts, when pleasure comes from movement toward a fake good. Therefore, in moral actions the conflict seems to occur more between different kinds of pleasures, which ultimately depends on different kinds of love (i.e. “natural love” and “self-love”), rather than between the sensible and rational dimensions of human beings. However, if this is accurate, we have to solve at least two problems. The first problem is that, according to Malebranche, sensibility concerns the body, which Malebranche considers as the source of every error.<sup>9</sup> How, then, can we state that morality has something *positive* to do with the sensible domain<sup>10</sup> without contradicting Malebranche’s claim regarding the body? The second, and related, problem concerns the relationship between “natural love” and “self-love”. How can they be in conflict if every movement of the will comes from God? I will attempt to answer these questions in the following sections.

## Good and Bad Pleasures

Malebranche’s critique on the (negative) role of the body in human life is present in the first lines of the *Recherche*. Here, Malebranche explains that human minds lie in a midway between the “Creator” and “corporeal creatures” (RDV, Preface, SAT XXXIII). Whereas

---

8 The discussion between Malebranche and Lamy is an episode of the debate on Quietism, a spiritual and mystical movement widespread in France in the second half of the seventeenth century, according to which the love of God must be completely disinterested to be meritorious. On Quietism, see Armogathe 1973. On the Malebranche-Lamy debate, see Priarolo 2018.

9 “The mind’s union with the body, on the contrary, infinitely debases man and is today the main cause of all his enon and miseries.” RDV, Preface, SAT XXXIII.

10 I have emphasized the word “positive” since without specifying it, Hume could also have agreed with it. In fact, we could say that morality has to do with sensibility in the sense that acting morally means *denying* the sensible dimension of human beings.

the union with God “raises the mind above all things” and provides the human mind with “its life, its light, and its entire felicity” (ibid.), the union with the body is the source of “all [our] errors and miseries” (ibid.). This condemnation of the body, which is present in the history of philosophy at least since Pythagoras,<sup>11</sup> comes to Malebranche from Augustine, one of the two main sources of Malebranche’s *Recherche*. However, as is well known, the other source is Descartes. Although Descartes’ philosophy owes a lot of debt to the Platonic tradition,<sup>12</sup> and both Plato and Descartes are diffident regarding the corporeal dimension’s role in the domain of knowledge, Descartes’ general attitude toward corporeity is more complex. Besides *Traité de l’homme* – the first book of Descartes read by Malebranche, and which, according to his biographers, “converted” him to philosophy in 1664<sup>13</sup> – in which Descartes explains all the things that the body-machine can do by itself, by giving to the corporeal mechanism a lot of value,<sup>14</sup> the *Meditations on First Philosophy* also proposes a positive account of sensibility. Although sensations cannot provide reliable knowledge of the world and, in this respect, are a great source of errors – they are “obscure and confused”<sup>15</sup> – their role is essential in ordinary life. Sensations, in fact, “inform the mind of what is beneficial or harmful for the composite of which the mind is a part” and with regard to this “they are sufficiently clear and distinct”.<sup>16</sup> Sensibility has, therefore, an adaptive function in human life, and despite some inconveniences that could occasionally happen,<sup>17</sup> “in matters regarding the well-being of the body” the senses “report the truth more frequently than not”.<sup>18</sup> Like Descartes, Malebranche emphasises that sensations play a fundamental role in our life, although they do not correctly describe, or precisely *because* they do not correctly describe the events. According to Malebranche, indeed it is more useful for us to feel pain in the hand we have placed too close to the fire than to know the precise corporeal modifications the fire is causing in it:

---

11 This is at least Plato’s reading of Pythagorean philosophy in the *Phaedo*: cf. Palmer 2014, 211.

12 On the extent of Plato’s presence in Descartes’ thought, see, for instance, Menn 1998 and – more cautious – Wilson 2008.

13 The story of this encounter is legendary: “while walking in the rue Saint-Jacques he [Malebranche] asked whether there were new books and he was offered Descartes’ *Treatise on Man* that M. Clerselier had just published [...]. He liked so much the method of reasoning and the mechanics which he read leafing through it that he bought the book and read it with so much pleasure that he was forced to interrupt his reading from time to time because of the heartbeats which arrived to him, so intense it was the pleasure of reading it.” OC XVIII, 46.

14 On the *Traité de l’homme* and its influence in post-Cartesian philosophy, see Scribano 2015.

15 See, for instance, Descartes, *Sixth Meditation*, in Descartes 1974–1989 (hereinafter AT followed by the volume number and page number) AT VII 80, English translation in Descartes 1984–1991 (hereinafter CSM followed by the volume number and page number) CSM II 55.

16 Descartes, *Sixth Meditation*, in AT VII 83, CSM II 57.

17 According to Descartes, on some rare occasions, for instance when we are sick, our body can give us information that is not useful but harmful for us, for example, drinking when we are in a state of dropsy. See Descartes, *Sixth Meditation*, AT VII 85–9 CSM II 59–61.

18 Descartes, *Sixth Meditation*, in AT VII 89 CSM II 61.

if the soul perceived only what takes place in the hand when it is being burned, if it saw in it only the movement and rupture of fibers, it would hardly take any notice; it might even derive from it some whimsical satisfaction, like those simpletons who amuse themselves by breaking everything in furious orgies of destruction. (RDV, I, 10, V, SAT 51)

Hence, sensations give us a sort of practical rather than epistemic information, and pain and pleasure are no exception:

while feeling pleasure and pain, which are things differing more than in degree, we more easily distinguish the objects that occasion them [...] if we must either embrace or flee the bodies surrounding us, this way of informing us is the quickest. (RDV, I, 10, V, SAT 52)<sup>19</sup>

For this reason, we cannot blame our senses and conclude that they are “corrupt and disordered” (RDV, I, 5, I, SAT 23).

For if it be considered that they are given us for the preservation of our body, it will be seen that they fulfill their purpose perfectly well, and that they conduct us in so faithful and appropriate a fashion to their end that it seems wrong to accuse them of being corrupt and disordered. (ibid.)

Therefore, the problem is not pleasure itself – we cannot “tell men that sensible pleasures are not good [...] because this is not true,” Malebranche claims (RDV, IV, 10, 1, SAT 309) – but the fact that we do not recognize the true source of pleasure: God. As we read in the *Traité sur la nature e la grace*<sup>20</sup>

[w]hen the soul loves only its own pleasure [...] it is unjust, it is ungrateful, it is blind, if it loves its pleasure, without offering the love and the respect that is due to the true cause which produces it in it. (TNG, III, IV, TG 171)

The real goodness of pleasure lies, then, in the fact that its source is God, whose action on our soul produces pleasure in us through objects which are occasional causes of our modifications.<sup>21</sup> Consequently, pleasure is good not only because it shows us the “quick-

---

19 For a general account of the role of pleasure in Malebranche’s theory of perception, see Shapiro 2018.

20 The translation quoted is Malebranche 1992. In the main text the references will be TG (= Treatise on Nature and Grace) followed by the page number.

21 See also RDV IV, X, I: “only God is sufficiently powerful to act in us, and to make us feel pleasure or pain. For it is obvious to every man who consults his reason and who scorns the reports of his senses that it is not the objects of sense that act effectively in us, since body cannot act upon mind; neither does our soul cause its pleasure and pain in itself upon the occasion of objects, for if the suffering of pain were dependent upon the soul, it would never suffer pain.” SAT 307–8.

est” way to survive, but also because its ultimate cause is God.<sup>22</sup> In this sense, we could say that the problem concerning pleasure is related to the ignorance of its real source and that Malebranche is claiming that, since we cannot but love what pleases us, if we do not understand that the real cause of pleasure is God, instead of loving God, we end up loving the bodies which we consider the source of pleasure. In other words, we could suppose that according to Malebranche, self-love can prevail over other inclinations and become the leading movement of the will only because we ignore the true good. However, if this is the case, does this not mean that in Malebranche’s account moral actions result from conflict between the rational dimension, which could or could not show us the right good, and the sentimental dimension, which blindly follows what appears good to the intellect and negatively contributes to true knowledge because of its intrinsic obscurity and attraction to pleasures?<sup>23</sup> In support of this interpretation, which is consistent with the above-mentioned remark of Hume on Malebranche’s “abstract theory of morals”, we could quote some passages in which Malebranche states, for instance, that to be meritorious love “has to arise or proceed from the light: it has to be regulated only by reason” (*Méditations chrétiennes*, OC X 154) – or as we read in the *Traité de Morale*, “whoever goes back into himself most deeply, and who listens to the inner truth in the greatest silence of the senses, imagination and passions, is most solidly virtuous” (*Traité de Morale*<sup>24</sup> – hereinafter TM – I, II, XII, 57).<sup>25</sup> In *Conversations chrétiennes* (1677), in a passage that Lamy considered a proof of Malebranche’s proximity to the doctrine of Quietism, Malebranche claims that in order to be meritorious, love must be *independent* from the pleasure that we could possibly obtain from the object of our love.

God is so lovable that who sees him as he really is would love him even in the middle of the greatest pains; and it is not to love him as he deserves, to love him only because he is the only one who can cause in us pleasant feelings. (*Conversations chrétiennes*, VIII, OC IV 179)

Is, then, Malebranche proposing a refusal of sensibility and a correspondent ethical doctrine, according to which all we have to do in order to behave properly is to leave the sensible dimension in each of its aspects and try to reach the spiritual, immaterial, intelligible

---

22 See TNG, III, IV: “The word good is equivocal: it can signify either the pleasure which makes one formally happy, or the cause of pleasure, real or apparent.” TG 171.

23 It is worthwhile recalling that according to Malebranche we know our mind “par sentiment intérieur”, i.e. through “inner sensation”, and that the main feature of this kind of knowledge is precisely obscurity: see RDV, III, II, VII, IV, SAT 237–9.

24 The translation quoted is Malebranche (1993). In the main text the references will be TE (= *Treatise on Ethics*) followed by the page number.

25 See also TM, II, V, III; “The Son sends the Holy Spirit, which precedes Him and the Father, though in a unity of principle; our love also presupposes the light, proceeds from it and is produced from it. Finally, love which proceeds from a clear understanding, is loved by itself and by the object of its knowledge and by the understanding itself.” TM 163–4.

realm? The answer is neither yes nor no. In fact, Malebranche appears divided between full adhesion to a Platonic-Augustinian critique of everything that pertains to corporeity and a more “naturalistic” reflection, a legacy of Descartes, on the mechanism of human behaviour. This tension runs throughout his writings and will be resolved *via* what has been called “sensible morality”,<sup>26</sup> an ethical theory that not only cannot overlook the sensible dimension of human beings, but also conceives reason, or better the rational Order, as something that must become sensible to motivate action. In understanding this point, it is important to stress the fact that Malebranche’s definition of the will as a general inclination toward good was replaced over time by the definition of the will as a desire to be happy. This shift increases the weight that pleasure has in motivating actions. As we read in the *Traité sur l’amour de Dieu*:

It is absolutely impossible to will something if nothing touches us: it is impossible that the soul is shaken, that it receives an impression, a movement, if nothing hits it. But there are different kinds of pleasure. Enlightened pleasure, bright, reasonable, which brings one to love the true cause of it, to love the true good, the spiritual good; confused pleasure, which excites love for powerless creatures, for fake goods, for corporeal goods. (OC XIV 9–10)

Rationality serves, then, to “enlighten” pleasure, but it is not rationality that moves the will. The driver is indeed pleasure, which is “by itself efficacious in relation to the will, for it moves and urges it, as it were, toward the object”<sup>27</sup> (RDV, Elucidation I, SAT 555). Consequently, to give the right direction to the will, the divine grace that Jesus Christ spreads on human beings must also be a “grace of feeling”, i.e. a delectation, a *pleasure*, which only can balance concupiscence.

It is certain that pleasure makes him who enjoys it happy, at least during the time that he enjoys it. Thus, men being made to be happy, pleasure always sets their will in motion, and puts it ceaselessly in motion towards the object which causes it or which seems to cause it [...]. Now concupiscence consists only in a continual chain of feelings and of movements which precede reason and which are not sub-

---

26 Bardout 2000, 111. According to Bardout, the cause of this change in Malebranche’s ethical theory is the theory of “efficacious ideas”. On this theory, see Robinet 1965, 300–51. Also, Ginette Dreyfus states that Malebranche’s ethics evolves in this sense. See Dreyfus 1958, especially p. 383. For a critique of Bardout, see Schmaltz 2007, 102–3.

27 The thesis, according to which the rational dimension does not move the will, can be found already in the RDV, V, X: “There is a great difference between the mind’s pure ideas and the soul’s sensations and emotions. The mind’s pure ideas are clear and distinct, but it is difficult to make them familiar. The soul’s sensations and emotions, on the other hand, are quite familiar, but it is impossible to know them clearly and distinctly. Extension, numbers, and their properties are known clearly; but unless they are made sensible by means of characters expressing them, it is difficult to represent them, for what is abstract does not affect us.” SAT 395.

ject to it [...]. It was, then, necessary that the second Adam [...] produces in us pleasures and horrors contrary to those of concupiscence [...]. Thus the grace of which Jesus Christ is the occasional cause [...] is not the grace of enlightenment [...] it is rather the grace of feeling. (TNG, II, XXXI, TG, 151)

According to Malebranche's mature ethics, therefore, a moral act is not an act that excludes pleasure, but an act that reflects on pleasure and, after this reflection, evaluates whether the pleasure is "enlightened" (i.e. licit and just) or "confused" (i.e. illicit and unjust). This evaluation is what Malebranche calls "consent", which is the approval of the determination of our will. To this regard,

every pleasure or material motive, although efficacious by itself in relation to the will it moves, is not efficacious by itself in relation to the will's consent; for it does not remove the soul's desire to be genuinely happy, or the power to withhold its consent and to examine whether such a pleasure accords with the sovereign happiness it invincibly desires. (RDV, Elucidation I, SAT 555)

Defining the moral act in terms of consent permits Malebranche to overcome problems concerning the doctrine of occasionalism, in which has fallen his first definition of freedom as the power to change the direction of the will.<sup>28</sup> According to this "new" definition, the will is *always* determined by pleasure. As a consequence, as the movement of bodies, it conserves not only the same quantity of movement but also the same direction, and therefore it does not interrupt the movement impressed in it by God.<sup>29</sup> The growing importance and development of Malebranche's occasionalism are essential to understanding this passage from a deep contempt of the sensible dimension to the recognition of its "neutrality", even though, as we have mentioned, Descartes' influence introduced a less negative approach to sensibility in a very critical context. As we will see, the question of the relationship between self- and natural loves which we evoked in the previous section, will meet a similar fate.

## Different Loves

The relationship between self-love and the love of God (i.e. natural love) is a very critical point in Malebranche's thought. As we have seen, self-love, like every other inclination, has its source in God, who loves himself and his works and wants to transfer this love to all creatures. "He loves Himself, He loves us, He loves all His creatures; therefore He creates

---

28 How could a finite being interrupt a movement begun by God himself? On Malebranche's problematic definition of freedom, see Kremer 2000, Schmaltz 1996.

29 In this sense, the debate with Leibniz on the laws of motion is very important in order to understand the changes that occurred in Malebranche's definition of freedom. See Priarolo forthcoming.

no mind without inclining it to love Him, to love itself, and to love all creatures.” (RDV, VI, 1, IV, SAT 267) Self-love has, therefore, the same origin and the same aim as the love of God and of creatures. However, in the first four editions of *Recherche*, Malebranche claims that original sin perverted the nature of self-love to the point that this kind of love became the opposite of the love of God.<sup>30</sup>

Nonetheless the inclination that we should have toward God has been lost because of the sin, and it has only left in our will an infinite capacity for all goods or for the good in general, and a strong inclination to possess them which can never be erased. But the inclination we should have toward our conservation, or our self-love, has so increased that it has finally become the absolute master of the will. It has also changed the love of God, or the inclination we have for the good in general, and the love we should have for other men in its nature. In fact, it can be said now that we have love but for ourselves, because we love everything but for us, whereas we should love only God and everything else for God. (RDV, IV, OC II, 45–6, my translation)

According to Malebranche, after original sin and without grace it is impossible for us to love anything unless we have an interest in doing so, and “we currently receive some pleasure by loving it” (RDV, IV, OC II, 47). Moreover, a similar position is presented in the passage of the *Tenth Elucidation* quoted above, in which Malebranche states that self-love “conceals Order from us and prevents us from following it” (SAT 619). In this sense, we can say that although the first occurrences of self-love in RDV include it among the natural inclinations, which deriving from God are necessarily positive, Malebranche’s first analysis of self-love, at least until 1678, attributes it a very negative role.

Conversely, in later writings such as *Traité sur l’amour de Dieu* (1697) this perspective appears completely changed.

The love of God, even the purest, is interested in this sense, because it is excited by the natural impression that we have for the perfection and the happiness of our being, in one word for the pleasure in general, or for the agreeable perceptions which are related to the true cause which produces them and which make us love it. (OC XIV, 23)

Far from being opposites, as in RDV, *Christian Conversations* and *Elucidations* to RDV, in *Treatise on the Love of God* self-love and the love of God are seen as complementary. Some commentators have supposed that this change is due to the previously mentioned debate on “pure love”, which obliged Malebranche to distance himself from the Quietist position.<sup>31</sup> However, it is important to note that the previous *Traité de morale* (1684) already

---

30 On Malebranche’s doctrine of original sin and its relationship with occasionalism, see Bozovic 1998.

31 For instance, Bardout 2000, but also Robinet 1965.



contained a similar position, and interpreted self-love in a positive way. Here, Malebranche explains that since it is a product of God, “self-love, by itself, is not bad” (TM I, VIII, XIV, TE, 104) and “is neither a virtue nor a vice” (TM II, XIV, IV, TE, 220). Quite the contrary, self-love is the source of every “motive” of our will, what gives the will “*du mouvement pour aller plus loin*” (RDV, I, I, II, OC I 48).<sup>32</sup>

Every motive is naturally and necessarily based on self-love, on our invincible desire to be happy – where I mean solidly happy. Every motive is based on the movement God constantly impresses upon us toward happiness and the perfection of our being. In a word, every motive is based on self-will, for we can only love by way of our own will. (TM I, VIII, XIV, TE, 104)

As I have shown elsewhere (cf. Priarolo 2018), this change – which, by the way, makes more consistent Malebranche’s first thesis, according to which self-love derives from God as every other inclination – is probably due to Arnauld’s criticism of Malebranche’s definition of freedom. How is it possible, Arnauld asked in *Des vraies et fausses idées* (cf. Arnauld 1986– orig. 1683–, 258), to state that the will can turn “this impression [toward the general good] toward objects that please us so that our natural inclinations are made to settle upon some particular object” (RDV, I, I, II, OC I 48, SAT 5) without contradicting occasionalism? Is Malebranche claiming that human will is a real power? After a deep and complex reflection that continues until his last work, *Réflexions sur la prémotion physique* (1715), Malebranche answered that we do not and cannot stop the movement of the will as he wrote in the first editions of the *Recherche*,<sup>33</sup> but we give or do not give our consent to the determination produced by pleasures. As we read in the passage of the *First Elucidation* already quoted:

Every pleasure or material motive, although efficacious by itself in relation to the will it moves, is not efficacious by itself in relation to the will’s consent; for it does not remove the soul’s desire to be genuinely happy, or the power to withhold its consent and to examine whether such a pleasure accords with the sovereign happiness it invincibly desires. (RDV, Elucidation I, SAT 555)

Far from only being related to self-love, pleasure is essential to every form of love because it is part of the mechanism of the will itself, which, as we have seen, is nothing but the desire to be happy:

---

32 The English translation is less suggestive: “The mind tends to proceed still further”, SAT 5.

33 “Our sin consists precisely in the fact that we stop at a particular good the impression that God gives us in order to love every good or the universal good during the time in which we have to and we can love it. Therefore the sin is nothing, and though God does everything, he does not do it.” OC III 24, III–IV editions (1678).



we are materially predetermined toward the good in general, because we necessarily will to be happy and because the desire for happiness is in us independently of us [...] we are also materially predetermined toward particular goods in this sense, that we are urged toward what we know and relish as good. The soul's natural impulse toward particular goods is, in effect, but a natural consequence of its impulse toward the good in general. Thus, all pleasure is by itself efficacious in relation to the will, for it moves and urges it, as it were, toward the object. (RDV, *First Elucidation*, SAT 555)

This implies that it is impossible to love something without any reference to this desire and to the pleasure we expect from its satisfaction. Consequently, as Malebranche writes in *Traité de l'amour de Dieu*, “[i]f we love the Order, it is because the beauty of the Order is pleasant; if we love the wine, it is because of the pleasure that we find in drinking” (OC XIV, 9). If virtue is the love for Order,<sup>34</sup> it is impossible to act morally without taking pleasure in it – i.e. without the involvement of our affective dimension. Reason is, of course, still essential for morality, because it is reason that can “enlighten” pleasures and show whether they are good or bad pleasures. Nonetheless, to reach the supreme end of human beings (i.e. union with God) the two dimensions of human beings, far from being opposites as Hume thought, must cooperate.

## Conclusions

Two additional points remain to be addressed. The first concerns the reason why Hume claims that Malebranche poses “an abstract theory of morals”, if, as we have seen, it is very difficult to attribute to Malebranche a similar opinion, because of the role that affectivity has in his ethics. It must be noted that Hume's knowledge of Malebranche's writings concerns mostly the *Recherche*.<sup>35</sup> But in the *Recherche* the definition of the relationship be-

---

34 “The Love for Order is not only the principal of the moral virtues; it is the unique virtue.” TM, II, I, TE, 53.

35 In a letter written to his friend Michael Ramsay on 31 August 1737, after Hume left the College La Flèche, we read: “I shall submit all my Performances to your Examination, and to make you enter into them more easily, I desire of you, if you have Leisure, to read once over le Recherche de la Verite of Pere Malebranche, the Principles of Human Knowledge by Dr Berkeley, some of the more metaphysical Articles of Bailes Dictionary; such as those [...] of Zeno, and Spinoza. Des Cartes Meditations would also be useful but don't know if you will find it easily among your Acquaintances. These Books will make you easily comprehend the metaphysical Parts of my Reasoning and as to the rest, they have so little Dependence on all former systems of Philosophy, that your natural Good Sense will afford you Light enough to judge of their Force and Solidity.” This letter was published for the first time in 1963 by Tadeusz Kozanecki in the *Archiwum Historii Filozofii i Myśli Społecznej*. I quote from Popkin 1964, 775. The library of the University of Edinburgh possesses Hume's copy of Malebranche's RDV. It is the Lyon edition of 1684. See [https://exhibitions.ed.ac.uk/record/23160?highlight=\\*.](https://exhibitions.ed.ac.uk/record/23160?highlight=*.)

tween the reason and the will, is actually conceived as a subordination of the latter to the former. As Malebranche writes in the *Recherche*:

it must be carefully noted that insofar as a mind is thrust toward good in general, it cannot direct its impulse toward a particular good unless that same mind, insofar as it is capable of ideas, has knowledge of that particular good. In plain language, I mean that the will is a blind power, which can proceed only toward things the understanding represents to it. As a result, the will can direct both the impression it has for good, and all its natural inclinations in various ways, only by ordering the understanding to represent to it some particular object. The power our soul has of directing its inclinations therefore necessarily contains the power of being able to convey the understanding toward the objects that please it. (RDV, I, I, II, OC I 48, SAT 5)

Nevertheless, as we have seen, the idea that the soul has a power, whatever it is, clashes with occasionalism and obliges Malebranche to modify his thoughts on the role of reason and pleasure in human actions. As a result, as Bardout remarks “[t]he will is no more defined as the basic movement toward the good in general, determined by clearly known relations of perfection. The indeterminate movement toward the good bends to a radical desire of happiness.” (Bardout 2000, 118). Hume was, therefore, right in attributing moral rationalism to Malebranche, if Malebranche had stopped at the first editions of the *Recherche*, but he was certainly wrong if we consider the evolution of Malebranche’s ethics.

The second point regards the role of Malebranche’s ethical thoughts on the sentimentalist approach to morality developed by eighteenth-century British thinkers, a topic that we can only briefly mention here. The presence of Malebranche’s thoughts beyond the English Channel is widely acknowledged,<sup>36</sup> but Malebranche’s influence has been considered mostly with regard to his theory of knowledge by dwelling upon who accepted it, as in the case of John Norris, or upon who refused it, as in the case of John Locke. Less known is the role that Malebranche’s other themes played in British philosophy. Among these is, in my opinion, the doctrine of inclinations. As previously mentioned, the origin of the tension between self- and natural love lies in the fact that, according to Malebranche, all inclinations derive from God’s impression on the human soul. It is also for this reason that Malebranche reworks the opposition between self- and natural love: in fact, since both of them have their source in God, it would be inconsistent to state – as Malebranche did in his first writings, but not at all in the later ones – that they are in conflict with each other. Now, Malebranche’s notion of inclination establishes the root of moral behaviour in some natural features of human beings. As we have seen, besides and because of the general inclination toward God, human beings are driven by two other main inclinations: self-love and the inclination towards others. Like the selfish or egoistic behaviours, the good attitude that we (possibly) have toward others is then based on the way in which human beings

---

36 See McCracken 1983.

are constituted and is rooted in human nature. A careful reader of Malebranche, Francis Hutcheson, in his *An Essay on the Nature and Conduct of the Passions and Affections*, in line with Malebranche's definition states that

our affections are contrived for good in the whole. Many of them indeed do not pursue the private good of the agent; nay, many of them, in various cases, seem to tend to his detriment, by concerning him violently in the fortunes of others, in their adversity, as well as their prosperity. But they all aim at good, either private or publick. and by them each particular agent is made, in a great measure, subservient to the good of the whole. Mankind are thus insensibly link'd together, and make one great system, by an invisible union [...]. Thus we are formed with a view to a general good end. (Hutcheson 2002, 118)<sup>37</sup>

For this reason, against moral egoists, such as Hobbes and especially Mandeville, Hutcheson states that moral actions are not "artificial"; that is, they do not derive from a rationalistic refusal of our natural motives, but flow from a feature, an "affection" in Hutcheson's words, that is as natural to us as self-interest – i.e. benevolence. In acting morally, therefore, we perceive a peculiar kind of pleasure, which demonstrates that what determines our moral evaluations is not reason, but a "moral sense":

some actions have to men an immediate goodness; or, [...] by a superior sense, which I call a moral one, we perceive pleasure in the contemplation of such actions in others, and are determin'd to love the agent, (and much more do we perceive pleasure in being conscious of having done such actions our selves) without any view of further natural advantage from them. (*An Inquiry concerning Moral Good and Evil*, Hutcheson 2004, 88)

In Hutcheson's writings, therefore, the idea that our moral behaviour does not depend on reason but is part of our experiencing life is consistent with a conception of the nature of human beings that sees our relationships with others as a fundamental part of it. Evoking Malebranche's considerations on human nature, Hutcheson then shows that moral actions are connected more with the sensible, instinctual dimension of human beings than with an abstract evaluation of what is good or bad. In this sense, we can observe that Malebranche's reflection on the role of inclinations shows that one of the effects of defining human ethical behaviour as starting from the affective constitution of human beings is that the role of reason ceases to be fundamental in defining morality. Consequently, although starting from a theological framework, Malebranche seems to attain a *naturalistic* view of human behaviour, a view that the moral sense theory will develop, and that will definitely triumph precisely with the philosophy of David Hume.

---

37 On the influence of Malebranche on Hutcheson's doctrine of the "calm desires", see Jensen 1971.

# Bibliography

## Primary Sources

- Aquinas, Thomas. 1984. *Summa theologiae*. Latin text of the leonine edition. 35 vols. Bologna: Edizioni studio domenicano.
- Arnauld, Antoine. 1986. *Des vraies et des fausses idées* (1683). Paris: Fayard.
- Descartes, René. 1974–1989. *Oeuvres de Descartes*. Edited by Charle Adam and Paul Tannery, 11 vols. Paris: Vrin.
- Descartes, René. 1984–1991. *The Philosophical Writings of Descartes*. Edited by John Cottingham, Robert Stootholf and Rupert Murdoch, 3 vols. Cambridge: Cambridge University Press.
- Hume, David. 1983. *An Enquiry Concerning the Principles of Morals*. Edited by Jerome B. Schneewind. Cambridge–Indianapolis: Hackett.
- Hutcheson, Francis. 2002. *An Essay on the Nature and Conduct of the Passions and Affections with Illustrations on the Moral Sense*. Edited by Aaron Garrett. Indianapolis: Liberty Fund.
- Hutcheson, Francis. 2004. “An Inquiry concerning Moral Good and Evil.” In *An Inquiry into the Original of our Ideas of Beauty and Virtue in Two Treatise*, edited by Wolfgang Leidhold. Indianapolis: Liberty Fund.
- Malebranche, Nicolas. 1958–1984. *Oeuvres complètes de Malebranche*. 20 tomes and 2 vols. Paris: Vrin-CNRS.
- Malebranche, Nicolas. 1992. *Treatise on Nature and Grace*. Edited by Patrick Riley. Oxford: Clarendon Press.
- Malebranche, Nicolas. 1993. *Treatise on Ethics*. Edited by Craig Walton. Dordrecht: Springer.
- Malebranche, Nicolas. 1997. *The Search after Truth*. Edited by Thomas M. Lennon and Paul J. Olscamp. Cambridge: Cambridge University Press.

## Secondary Literature

- Armogathe, Jean-Robert. 1973. *Le Quietisme*. Paris: PUF.
- Bardout, Jean-Cristophe. 2000. *La vertu de la philosophie. Essai sur la morale de Malebranche*. Hildesheim – Zurich – New York: Georg Olms Verlag.
- Bozovic, Miran. 1998. “Malebranche’s Occasionalism, or, Philosophy in the Garden of Eden.” In *Cogito and the Unconscious*, edited by Slavoj Zizek, 149–74. Duhram–London: Duke University Press.
- Dreyfus, Ginette. 1958. *La volonté selon Malebranche*. Paris: Vrin.
- Favaretti Camposampiero, Matteo, Priarolo, Mariangela and Scribano, Emanuela. 2018. *Occasionalism. From Metaphysics to Science*. Turnhout: Brepols.
- Jensen, Henning. 1971. *Motivation and the Moral Sense in Francis Hutcheson’s Ethical Theory*. The Hague: Martinus Nijhoff.

- Lee, Sukjae. 2008. "Occasionalism." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. <https://plato.stanford.edu/entries/occasionalism/>
- Kremer, Elmar. 2000. "Malebranche on Human Freedom." In *The Cambridge Companion to Malebranche*, edited by Steven Nadler, 190–219. Cambridge: Cambridge University Press.
- McCracken, Charles. 1983. *Malebranche and British Philosophy*. Oxford: Clarendon Press.
- Menn Stephen. 2008. *Descartes and Augustine*. Cambridge: Cambridge University Press.
- Moriarty, Michael. 2006. *Fallen Natures, Fallen Selves. Early Modern French Thought II*. Oxford: Oxford University Press.
- Nadler, Steven (ed.). 2010. *Occasionalism. Causation among the Cartesians*. Oxford: Oxford University Press.
- Nadler, Steven (ed.). 1993. *Causation in Early Modern Philosophy: Cartesianism, Occasionalism, and Preestablished Harmony*. University Park: The Pennsylvania State University Press.
- Palmer, John. 2014. "The Pythagoreans and Plato." In *A History of Pythagoreanism*, edited by Carl A. Huffman, 204–26. Cambridge: Cambridge University Press.
- Perler, Dominik, and Rudolph, Ulrich. 2000. *Occasionalismus. Theorien der Kausalität im arabisch-islamischen und im europäischen Denken*. Göttingen: Vandenhoeck und Ruprecht.
- Popkin, Richard. 1964. "So, Hume did read Berkeley." *The Journal of Philosophy* 61(24): 773–8.
- Priarolo, Mariangela. forthcoming. Amor che move. *Volonté et 'physique de l'âme' dans la Recherche de la vérité*. In *Nouvelles recherches sur la "Recherche de la vérité"*, edited by Jean-Cristophe Bardout, Denis Moreau and Vincent Carraud. Paris: Vrin.
- Priarolo, Mariangela. 2018. *Un plaisir divin: Malebranche, Lamy (et les autres)*. In *Le malbranchisme à l'épreuve de ses amis et de ses ennemis*, edited by Elena Muceni, Maria Cristina Pitassi, 149–68. Paris: Honoré Champion.
- Priarolo, Mariangela. 2017. *Universals and Individuals in Malebranche's Philosophy*. In *The Problem of Universals in Early Modern Philosophy*, edited by Stefano Di Bella and Tad Schmaltz, 142–65. Oxford: Oxford University Press.
- Riley, Patrick. 2003. *Malebranche and Natural Law*. In *Early Modern Natural Law Theories. Contexts and Strategies in the Early Enlightenment*, edited by Tim J. Hochstrasser and Peter Schröder, 53–87. Dordrecht: Kluwer.
- Robinet, André. 1965. *Système et existence dans l'oeuvre de Malebranche*. Paris: Vrin.
- Robinet, André. 1979. *Ordo/ordine dans l'oeuvre de Malebranche*. In *Ordo. II Colloquio Internazionale del Lessico Intellettuale Europeo*, edited by M. Fattori and M. Bianchi, 347–424. Rome: Edizioni dell'Ateneo e Bizzarri.
- Schmaltz, Tad. 1996. *Malebranche's Theory of the Soul. A Cartesian Interpretation*. Oxford: Oxford University Press.
- Schmaltz, Tad. 2007. *Malebranche on Natural and Free Loves*. In *The Concept of Love in 17<sup>th</sup> and 18<sup>th</sup>-Century Philosophy*, edited by Gabor Boros, Herman De Dijn and Martin Moors, 95–111. Leuven: Leuven University Press.

- Scribano, Emanuela. 2015. *Le macchine della mente. Fisiologia e metafisica tra Cartesio e Spinoza*. Roma–Bari: Laterza.
- Shapiro, Lisa. 2018. *Malebranche on Pleasure and Awareness in Sensory Perception*. In *Pleasure. A History*, edited by Lisa Shapiro, 124–45. Oxford: Oxford University Press.
- Vieillard Baron, J. 1996. “L’âme et l’amour selon Malebranche.” *Les Études philosophiques* (4): 453–72.
- Wilson, Catherine. 2008. “Soul, Body and World: Plato’s *Timaeus* and Descartes’ *Meditations*.” In *Platonism at the Origin of Modernity. Studies on Platonism and Early Modern Philosophy*, edited by Douglas Hedley and Sarah Hutton, 177–91. Dordrecht: Springer.

# Chinese Cultures of Love: The “Cult of *Qing*” and *Qing*’s “Naturalness”

## Introduction

Every civilization develops its own culture or cultures of love that manifest elaborations in controlling, interpreting and sublimating the libido and interpersonal relations. This paper will try to briefly introduce the way love was represented in philosophical and literary production of late imperial China, focusing on the so called “cult of *qing*”. *Qing* is a polysemous term, that can be rendered with various meanings, from “emotions” to “desires”, from “love” to “authenticity”. Thus we can call it “cult of *qing*”, “cult of love” or “cult of passions”. Chinese authors exploited the ambiguities of the *qing* term in various measures, playing with its relations with the concepts of “principle” (*li*) and “human nature” (*xing*), in the concurrent effort to rethink the relation between the affective and moral spheres. This new trend started in the middle of the Ming dynasty, under the influence of the Wang Yangming School of the Mind, *xinxue*, especially the *Taizhou School*, and its effects basically remained under the new Qing dynasty. Its main aim was the rehabilitation of emotions against the orthodox rigorist contrast between moral principles and desires in the ambit of Neo-Confucianism.<sup>1</sup> When it came to the task of re-evaluating *qing*, scholars have attributed a variety of meanings and values to it, from a “domesticated and moralised sentiment” to the elaboration of individual spontaneous and genuine aspirations. In this process several writers reached various compromises in the negotiation between the two

---

1 By “Neo-Confucianism” I mean the Confucian revival, which started in the Song dynasty (960–1279) and lasted throughout the entire second millennium. The term has been disputed for its vagueness, and other names have been proposed in its place. Here I will disregard the terminological issue and maintain the term for its practical utility. Considering the entire flow of thought that refers to the Confucian themes in the course of the second millennium, it does not contradict the pluralism of intellectual life in China in the same period.

extremes, the rigorous endorsement of heavenly principles and the strong support of love passion and desires.

Although the term “cult of *qing*” is a label easily misused, it is a pregnant concept for understanding the changes in the impact of Confucian morality on Ming and Qing society, the intellectuals’ attitudes developed under the loosening of political control during the decline of Ming dynasty, the further progress of urbanisation and commercialisation, and the rise of a merchant class. This new sensibility was not only manifested in the search for novel fashions and styles among writers, it also influenced intellectual élites of society and the growth of a kind of cult of self. The rise of a cult of passions, with its various forms, the sublimation of love, the crisis of traditional gender roles and the increasing space given to the self by writers reflect the evolution in the Chinese imaginary world in late imperial China. The “cult of *qing*” marks a new trend in philosophical and moral fields, as well as in literary production, by re-discussing the role and evaluation of emotions and desires. By positing the centrality of emotions, challenging the priority of principle regarding feelings and desires, and declaring their universality, it proposed a different vision of man and morality. *Qing* is a key concept for understanding the intellectual life and its polyphonic contributions.

## Talking of Love in Chinese Literature

“What is Love? A God, a Celestial Spirit, a state of mind?” was the question raised by Plotinus,<sup>2</sup> and many are the answers given in the West. With a different religious approach, “what is so powerful in the passion of love?” was the question Tang Xianzu 湯顯祖 (1550–1617) intended to experiment with in his drama “Peony Pavilion”. In the Preface he writes:

We do not know what originates passion, but, once it starts, it runs so deep: those who are alive can die, and those who are dead can be reborn by it.<sup>3</sup>

In a famous novel of the eighteenth century, the “Dream of the Red Mansion” (*Honglou meng* 紅樓夢), the prelude of a prophetic opera held in the Land of Illusion<sup>4</sup> sibilantly retorts:

---

2 See the paragraph “Plotinus’ cosmic and psychological explications on love (*Erós*)” in the part written by Gábor Boros in Santangelo and Boros 2019.

3 “情不知所起，一往而深，生者可以死，死者可以生。生而不可與死，死而不可復生者，皆非情之至也”。*Tang Xianzu ji* [“Works by Tang Xianzu”] (*Mudan ting*, Tiji 題記), juan 2, 1094; transl. by Hsia 1970, 276.

4 The Land of Illusion is a *locus amoenus* of the “Dream of the Red Mansion”, which marks a crucial oneiric experience in the main character’s life, but is also a fundamental passage in the development of the novel.



When first the world from chaos rose,  
Tell me, how did love begin?  
The wind and moonlight<sup>5</sup> first did love compose.<sup>6</sup>

In 1205, during the Jin dynasty, Yuan Haowen 元好問 (1190–1257) was going to Bingzhou in order to take part to the civil service examinations. En route, he met a hunter of wild geese, who told him:

Today I captured a goose and killed it. The other goose who escaped the nets, groaned and lamented, and could not leave; finally, it threw itself to the ground dead. Therefore, I bought them, and buried them in the bank of the Fenshui river, piling up some stones for memory, with the inscription of ‘Geese mound’. Then I wrote down the ‘Song of the Geese Mound’: ‘What is love in this world? Just to teach human beings that life and death are reciprocally related. 問世間情爲何物。直教人生死相許’. (Yuan Haowen, *Yanqiu ci* 雁丘詞)

The passage just quoted is famous in China, and in our modern sensitivity is read in connection with ecology and respect for nature. It is, however, a kind of allegory on love and death. It demonstrates how the question “What is love? What is the love passion?” was aesthetically raised all throughout the history of Chinese literature. In this poem, the author combines the dyad of death and love in a heroic but not dramatic way: the love of the geese unites them in life and death, and the religious ceremony of giving a grave to both together with writing of the epitaph are all expression of a genuine cult of love. With this broad meaning, we can say that this is one of many examples which witness the existence of an aesthetic cult of love before the Ming dynasty (1368–1644). In the Chinese literary production several documents extol love and present some elements that would be exploited by a kind of cult of love in any period, especially in the Chinese Middle Age, in the Tang and in the Yuan dynasties.<sup>7</sup>

As far as the concepts related to “love” are expressed in Chinese culture, the first impression is that the term *qing* – that traditionally implies an analogous, even if not superimposable broad semantic content – is not only a polysemous word, but it takes on different meanings and values according to author, context and period. Even if we concentrate the scope of enquiry on the seventeenth and eighteenth centuries, it ranges from the multifarious presentation made of it by the writer and publisher Feng Menglong 馮夢龍 (1574–

5 The expression “Wind and moon”, *fengyue* 風月 (lit. gentle breeze and bright moonlight), is one of the most common Chinese euphemisms for “love affair”.

6 See: Cao and Gao 1998, juan 5, 89; Hawkes’ transl. 1973, in Hawkes and Minford 1973–1986, I, 139.

7 For some precedents in Chinese literature and the role of *qing* in late imperial literature, see Hsieh 2007. On the specific cult of *qing*, just to mention a few studies, see Honda 1983; Wang 1999; Wang 1994; Li 1993, 1995; Chaves 1985; Huang 1998, 2001; Hanan 1981, 79–80, 95–97; Sun Chang 1990, 3–9; Epstein 2001, 61–119; Santangelo 2000b, 439–99; id [Shi Hualuo 史華羅] 2000a, 41–71; id, 2006, 133–63.

1646) to the tension between the destructiveness of desire and the naturalness of passion in the renowned novel “Dream of the Red Mansion”, from the magic subjective power in Pu Songling 蒲松齡 (1640–1715) to its identification with the Confucian virtues in Jiang Shiquan 蔣士銓 (1725–1784). Feng Menglong in his “History of Love” (*Qingshi*) attributed an almost mystic value to love, the great dramatist Tang Xianzu expanded the action of love after death infusing it with a new transcendent as well as narcissistic spirit, while Yuan Hongdao 袁宏道 (1568–1610) acknowledged the identity between desires-emotions (*qing*) and human nature (*xing* 性). Tang Xianzu’s dramas were so influential that their characters became models for many readers and spectators. This idealisation was carried on during Qing dynasty (1644–1911) by several writers – even if in a more problematic way.<sup>8</sup> In the “Dream of the Red Mansion” the search for an “impossible love”, pursues an ideal imbued with melancholy, and based on the rejection of social conventions and lack of faith in the human condition itself; doomed to encounter difficulty and misunderstanding in social relations, this love extends its interior dimensions beyond all measure precisely with the help of the world of imagination. And it is imagination that allows Pu Songling to reconcile desire and principle, defusing the tension and danger inherent in the unbounded nature of desire itself. And again it imbues Li Yu’s 李漁 (1611–1680) sensibility in his creation of a magic atmosphere at the presence of female charms. With extreme simplification, these concepts could be interpreted in various nuances between two essentially different but not exclusionary ways: the acknowledgement of the supremacy of love, or else its subordination to heavenly principles and social roles, in the reassertion of “moralised sentiments”. Not only various and contradictory concepts of *qing* circulated in late imperial China, but also new and old attitudes were blended, such as social stability and personal experience, aesthetics of sentiment and didacticism, the incorporation of the conventional values of chastity, friendship and filial piety in the cult of passions.

## Related Philosophical Concepts

For Chinese thinkers the concept of *xin* (心), “mind-heart” encompasses both the intellectual faculties – which the Western tradition attributes to the soul or to mind-reason – and the affective or sensory faculties – which the Western tradition normally attributes respectively to the heart and the senses. The character *xin* in the philosophical-religious field can be rendered with “mind-heart”, the principal seat of the emotions as well as of intellectual functions, although a few emotional qualities, such as courage, are located in other body organs. Moreover, in Chinese civilisation different dichotomies concern the psychological and moral representations, such as quiet and motion, heavenly principle and

8 Paul Ropp (2006, 203–28) has discussed the continuities and discontinuities in the cult of *qing* from the late Ming to the early nineteenth century, and the survival of the cult of *qing* long after the Manchu conquest, albeit in a somewhat sobered and sublimated form.

selfish desires, human nature (*xing*) and psychophysical energy (*qi* 氣), “moral mind” and “human mind”.<sup>9</sup> There is no enlightened and dispassionate reason ordering the wild and unbalanced energies of the irrational, and keeping a rein on them; it is rather the single mind-heart unit which sometimes retains its original equilibrium, reacting in a composed and timely manner and thus coinciding with the original nature of man and the universe, or may allow itself to be dragged along by the troubles of the world. Chinese thought never arrived at an opposition between the “light” of reason and the wild energies of the irrational in that the fundamental issue is always regarded as the pursuit of internal and external equilibrium, composure of the mind-heart seen in its fundamental oneness, conformity between stimulus and reaction. Summing up the main Chinese terms used in this article, *qing* 情, can often be rendered with “passions”, the inner response to external phenomena, and in other cases more specifically with love. Originally the notion of *qing* was considered very close to that of human nature, *xing*, a kind of its manifestation. Human nature in the Neo-Confucian system is the morality which is innate in human beings, man’s potential predisposition to achieving correct emotional reactions and behaviour. “Heavenly principle” (*tianli* 天理) is the “universal” present in every being, which is identified with “human nature” (*xing* 性) when it is “hung” (*gua* 挂) to human beings. The character *li* is usually translated as “principle” and sometimes – with danger of misinterpretation – as “reason” or “norm”. Chad Hansen proposes to render it with “natural tendency”, given the descriptive rather than normative character of Chinese thought (Hansen 1972, 183). In the course of the long evolution of *li*, influenced also by Buddhist debates, this term became one of the linchpins of Neo-Confucian doctrine and may be understood as the universal element present in every being. Principle is not “reason” in the sense it is used in Western thought, but rather the metaphysical personification of moral and natural order, which is embodied within the unique reality of the evolution of cosmic energy. The latter concept, *qi*, is rather the dynamic substance that fills and constitutes the universe, including objects and material energy as well as the vital force, and can be rendered as “cosmic and psychophysical energy” (whether we focus on the universe or the microcosm of body and man’s affective tendencies). The key point at the basis of moral judgement seems to lie in the fundamental distinction between internal and external harmony on the one hand, and psychological and social disturbance on the other, rather than in the degree of conformity to given rules or contents.

Acknowledging that the evolution and competition of different representations of *qing* is actually a very complex one and difficult to simplify, we can notice the progressive trend from an orthodox contrast between principles and desires (the metaphysical moral sphere and the dangerous emotional sphere) to a moderate re-evaluation of emotions, passions

---

9 In Neo-Confucian philosophy, the seat of man’s intellectual and affective capacities and the organ that partakes of the original goodness of human nature (*xing* 性) is the mind-heart, *xin*. Neo-Confucian thinkers used to distinguish the moral mind-heart, *daoxin* 道心 – which was identified with the heavenly principle – from human mind-heart (*renxin* 人心) that corresponds to the emotions in response to external impulses – that may be correct or not.

and desires, until the evocation of a kind of eroticism in which gender is rendered a floating performance. The so-called cult of *qing* was the expression of this new course started in the middle of the Ming dynasty, whose effects basically remained under the new Qing dynasty. In this process several writers reached various compromises in the negotiation between the two extremes, the rigorous endorsement of heavenly principles and the strong support of love passion and desires. When it came to the task of re-evaluating *qing*, scholars – such as Yang Shen 楊慎 (1488–1559), Luo Qinchun 羅欽順 (1465–1547), Yuan Huang 袁黃 (1533–1606), Tang Xianzu, Feng Menglong, Wang Fuzhi 王夫之 (1619–1692) – have attributed a variety of meanings and values to it, from a “domesticated and moralised sentiment” to the elaboration of individual spontaneous and genuine aspirations.

## Ideological Foundations of the Cult of *qing*

The effectiveness of imperial Neo-Confucianism, notwithstanding its austere morality and the puritan interpretation of some thinkers, lies precisely in the fact that it recognised man’s emotional reality and acknowledged its “naturalness”. In some way this followed the legacy of the ancient identification of feelings (*qing*) with the manifestation of human nature (*xing* 性). In practical terms it sought to reduce the scope of eros to the minimum and to channel its energies into the context of social requirements. Although in several passages love was extolled as the cohesive force of the universe, in many other cases its dark and sad aspects were emphasised. In the period from the sixteenth to the eighteenth centuries Chinese society was still dominated by Neo-Confucian orthodoxy, but the rigidity of morality created contradictions between individual aspirations and social rules in a society of great developments and changes. These contradictions were evident above all in the literary works. The Neo-Confucian dualism between principle and passions – and internal debate between supporters and challengers – is not a contraposition between functions of the same subject, but between different levels, one metaphysical-ethical and the other empirical. Parallel was the contrast between “excess” and “moderation” that might follow the rhetoric of the “egoism-universalism” dyad (*si-gong* 私公), in which the second term not only indicates what is public, but also what is objective, not bent to the particular, and in harmony with the world. The dangers of excess were commonly the subject of worry, and writers attributed the moral responsibility to the so-called “four vices” (lit. yearnings), indulgence in wine, lust, greed and wrath.

In the late Ming period, however, new ways to understand the polarity of the “egoism-universalism” dyad emerged. Confucian orthodoxy prevalently maintained a moralistic attitude against what was considered the private and egoistic pursuit of profit, lust and greed (*si* 私, *li* 利, *tan* 貪) of immoral people.<sup>10</sup> This does not mean that Confucian tradition

---

10 In the West, compare the three corresponding appetites of drinking, loving and eating (*potoi*, *aphrodisia*, *edodai*) mentioned in Plato (*Laws*, III, VI, IX), in Aristotle (*Nicomachean Ethics*, III, VII) and

lacks a concrete recognition of the importance of the primary needs of individuals. A clear distinction was, however, stressed between need and desire, vital function and excess: food and procreation were legitimate needs, based on social requirements, and while excess was discouraged, Zhu Xi distinguished between gluttony and hunger.<sup>11</sup> A radical change can be noticed, for instance, with Lu Ji's 陸楫 (1515–1552) discourse on luxury, that elaborated an economic analysis which corresponded to the standard of living and to the advanced economic conditions of the Jiangnan region (Lu 1937, 3–4, cit. in Xu 1986, 103). His arguments justified the pursuit of pleasure, such as ostentatious visits to Western Lake, for the positive effects created by the trade of luxury objects, including prosperity, wealth and employment. This perspective in economy corresponded to a new attention paid toward individual interests and desires. Luo Qinchun 羅欽順 (1465–1547) states:

The fact that man has desires definitely derives from heaven. Some are necessary and cannot be repressed; some are appropriate and cannot be changed. If those that are irrepressible all conform to the principle of what is appropriate, how can they not be good? It is only heedlessly giving way to passions, indulging the desires, and not knowing how to turn back that is evil. Confucians of the past often spoke about eliminating or restraining human desires and thought it necessary to resort to severe means in order to repress them. But their mode of expression seems one-sided and exaggerated. The desires, together with pleasure, anger, sorrow and joy, are qualities of nature. Can pleasure, sorrow, anger, and joy also be eliminated? (Luo 1999. Cf. also Huang Zongxi 1984, juan 47, 490–2)

Scholars (for instance Huang 1998, 2001; Wang 1994) trace back the ideological basis for the late Ming cult of *qing* to Wang Yangming 王陽明 (1472–1529). The centrality of the mind-heart, its evaluation and conscience, in contrast to the “objectivity” of the rules settled by the orthodoxy, undermined the principle of authority regarding self-cultivation, thereby threatening one of the foundations of orthodox Neo-Confucianism. The reduction and transformation of the role of principles allowed more space and legitimation for emotions and feelings. According to some thinkers of the so-called Taizhou School – a branch of Wang Yangming's learning of the mind – like Wang Ji 王畿 (1498–1583), Wang Gen 王艮 (1483–1541), He Xinyin 何心隱 (1517–1579) and Li Zhi 李贄 (1527–1602), the self consists first of all of natural needs; they are physical as well as “selfish” inclinations, and are considered the motivations for self-improvement. Wang Gen's ideas of individual autonomy, self-respect, self-love, self-interest and selfpreservation can be summarised in the sentence “Learning of this innate conscience consists only of protection of the individual”

---

in Xenophon (*Memorabilia*, I), with the relative recommendations to avoid excess. Throughout the history of European thinking it is possible to perceive a clash between “real” and “false” needs, “good” needs and “bad” ones, and between “needs” and “desires.” Cf. Heller 1985, 314–7.

11 Zhu Xi's famous sentence distinguishes between eating for survival (*need*) and gluttony (*desire*): “飲食者，天理也；要求美味，人欲也” (Zhu 1983, juan, 224).

率此良知之學,保身而已, because the protection of the world is a consequence of the safety and dignity of the individual (吾身保, 然后能保天下).<sup>12</sup> These ideas differ from theories elaborated by Western thinkers, as the social being is projected to self-perfection, a microcosm which is part of the macrocosm. Self is seen as the root and society as the branch, and social harmony depends on the defence of the person.<sup>13</sup> For He Xinyin self-fulfilment meant the fulfilment of the needs of others, as the desires of individual people are universalised.<sup>14</sup> Either way, the individualistic discourse developed in the ambit of the Taizhou School was generally not in contrast with the contemporary social organisation. Their idea of *qing* was innovative and yet not contrasted with moral principles, but was considered the source of morality – the best means of consolidating human bonds (Lee 2007, 38).

Li Zhi extends He Xinyin's discourse on desires and does not stop short of recognising the basic vital needs. Li Zhi's recognition of desire as being positive in nature and his paradoxical praising of "selfishness" indeed represent a conscious change in the perception of individuality and society (Li Zhi 1959, "Deye ruchen houlu", juan 32, 544; Li Zhi 1975, "Shuda", "Da Deng Mingfu", juan 1, 40, and "Da Deng Shiyang", juan 1, 4). He expands this category to positive desires and goes on to praise "selfishness" as the starting point of any social process.

The wise have taught that 'one succeeds by not being troubled [overdoing things]' (*wuwei er cheng zhe* 無為而成者). But now the expression 'not being troubled' has come to mean 'absence of mind-heart' (*wuxin* 無心) [...]. If he does not use his mind-heart, the peasant leaves the field untilled, the craftsman works shoddily, and the scholar is not up to his job. How can something be obtained without the use of the mind? Commentators assure us that the so-called 'absence of mind' is

12 Wang 1975, *Wang Gen chidu* 王艮尺牘 (Wang Gen's Epistolary) 49: Yu Xu Zizhi 與徐子直 (Chinese Text Project, <https://ctext.org/wiki.pl?if=en&chapter=882061>). The basic concepts are those of respect, cultivation, protection of the *person* or of *self* (安身, 養神, 保身, 修身, 尊身, 守身, 敬身, 愛身, 反求諸起身), correlated with others' needs, desires and wishes (cf. Huang Zhouyue 2012, 59), in a reinterpretation of the Great Learning's (大學, 2) progressive process from self-cultivation to the world order. For Wang Gen the main passages start from the moral and especially personal care of self up to the empire (see *Mingru xue'an* juan 32, 315 and 317). On the difference between Confucian self-reform 修身 and Wang Gen's individual welfare, see Ruan 阮景东, 2014.

13 A comparison would be interesting between this concept of "self-preservation" (保身) and apparently analogous concepts in Europe, such as Pufendorf's "natural inclination of self-preservation" (see the paper presented by Heikki Haara at the Third Budapest Seminar in Early Modern Philosophy, October 2018). It is evident, however, how the context is different: in Europe it is related to debates on the normative foundations of natural rights, separation of morality and nature, and the presupposed links of inclination with God and/or rationality. In China, in contrast, naturalness and self-protection are constructed on the discussion about the relation between human nature (in its various interpretations) and emotions-desires, principles.

14 He 1981, 229; He Xinyin, *He Xinyin ji*, juan 2, 40–41, juan 3, 72, cit. in Jiang and Zhu 1989, 277. Cf. *Mencius*, 1b, 5.



not a real lack of mind, but only an absence of selfish feelings (*wu sixin* 無私心). Yet the mind of man is characterised by ‘selfishness’. (Li Zhi 1959, “Deye ruchen houlun”, juan 32, 544)

[T]he human mind is selfishness: Without selfishness, there is no mind 人必有私, 而後其心乃見; 若無私, 則無心矣. He who works the land must have an interest in the harvest so that he works diligently; he or she who works at home must have an interest in accumulating wealth in order to shrewdly manage the household affairs; the student must have an interest in passing his exams in order to study well. Likewise, if a functionary was not interested in receiving his pay, he would not have sought the job [...]. Desire for goods, for sexual satisfaction, for education, for a personal career, for the amassing of wealth; the desire to buy land or houses for one’s descendants; the request for geomantic protection to bring good luck to one’s children – all of these things are productive and keep life going in this world, all of these things are loved and practised by the people, all of these things are known and openly discussed [...]. The concrete principles of human relations consist of clothing and food. Without these items there would be no relationships [...]. Scholars should recognise the true essence that lies behind these relationships, and stop quibbling over them [...]. (Li Zhi 1975, “Shuda”, “Da Deng Shiyang”, juan 1, 4)

His paradoxical exaltation of “selfishness” was not a purely ethical or cynical discourse, but was the premise for a radical transformation in the perception of the individual and society. This means a different conception of human character, as personal identity is no longer related to the “moral self” according to the Classics, but to a “natural self,” supposedly free from ideological incrustations. In Li Zhi’s viewpoint, human nature is saturated with child-like qualities which, combined together, describe a genuine state of the mind (*zhenxin* 真心), the “child-mind” (*tongxin* 童心).<sup>15</sup> “Natural Self” corresponds to genuine human nature,<sup>16</sup> the child-mind, neither good nor bad, and to the individual’s desires: appetite for food, attraction for beauty, shelter from the cold, desire for elegance, quest for comfort, the fulfilment of one’s ambitions: “The pursuit of profit and the evasion of damage correspond to the common feelings of human beings.” 趨利避害, 人人同心 (Li Zhi

15 “Child-mind”, *tongxin*, is the term used by Li Zhi to indicate the original mind-heart before the influences of culture and education.

16 The concept of “genuine” is often linked to that of “authenticity” and “natural”, but it is used with innumerable meanings. For a modern psychological perspective it corresponds to the feeling that one is currently in alignment with one’s true or real self (Sedikides et al 2017, 521–5, online [https://www.researchgate.net/publication/316862597\\_State\\_Authenticity](https://www.researchgate.net/publication/316862597_State_Authenticity); concerning experimental research on the subjective criteria of evaluation, cf. Jongman-Sereno and Leary 2018. 1–32). In Europe this notion has been used in philosophy and aesthetics in different ways (for instance Kierkegaard and Nietzsche, the debate on the true and false self from Erich Fromm to Michel Foucault, not to mention the long tradition on natural law). In Chinese culture the basic concepts are based on Confucianism, Daoism and Buddhism.

1975, “Da Deng Mingfu”, juan 1, 41). This new concept of self and of the individual is not influenced by Western doctrines, but is the extreme result of a process in which traditional elements in Chinese thought were combined: the Daoist care for the body, which is considered part of the cosmos, the Buddhist freedom from conventions and rules, and the syncretism of non-orthodox elements concerning the innateness of emotions and desires (Santangelo 2016; 2017, 357–432).

Thus, the so-called “cult of *qing*” that grew at the end of the Ming dynasty developed under the inspiration of Buddhist “emptiness”, Daoist “permanent change” and “body’s care”, Neo-Daoist *xing-qing* proximity and rejection of conventional,<sup>17</sup> Confucian social morality, popular belief in retribution and predestination, instead of Western Neo-Platonism, Christianity or Sufism, or the metaphysical contrast between reason and passion, sin and the yearning for the Absolute, purity and angel-like idealisation of woman. It was nourished by a number of traditional myths and popular legends elaborated within the context of Chinese literature, such as the metamorphosis of flowers and fox-spirits, or the legendary figures of Zhuo Wenjun, Jiaoniang, Yingying and Liniang,<sup>18</sup> that played a role similar to the one played in Europe by the classical myths of Narcissus, Pygmalion, Psyche and Eros, or Venus or heroes such as Paolo and Francesca. On the other hand, the love ideal too draws upon its own allegories, such as the union of two trees through their branches and roots, the pair of birds, dreams, mirrors, flowers.

It was also thanks to literary contributions that in China the concept of love was gradually enriched over the centuries in both the conceptual-symbolic and the mythical levels. Just as the awareness of the importance of “talking of *qing*”, *jiangqing* 講情, can be traced back to *Mudan ting* 牡丹亭 and the debate surrounding it,<sup>19</sup> some aspects of this cult are evident in “The History of Love”, attributed to Feng Menglong. Here a few fundamental concepts related to *qing* are illustrated: *qing* is taken as the measure of all things, as being transcendent, common to all men, spirits, gods and all other beings, even inanimate things;<sup>20</sup> Feng stresses that it brings the moral requisites of Confucian virtue into line with

---

17 Cf. Epstein 2001, 304. Cf. also Zhou 2000, 339–74 who deals with the influences of Neo-Daoism on the School of Principle.

18 With the exception of Zhuo Wenjun, all are characters of famous literary works. See footnote 247.

19 See “*Mudan ting jiangqing* 牡丹亭講情” and “*Wang Jizhong bidian Mudan ting tici*” 王季重批點牡丹亭題詞, in Xu Fuming 1987, 42–3, 69, which stresses the opposition between talking of *qing* and talking of human nature. See also Li Wai-ye 1993, 60–2.

20 On love that transcends life and death, in the field of literature see Feng Menglong’s (1986b) *Qingshi*, juan 10, 308; on inanimate things juan 11, 320; chapter 20 deals with ghosts, chapter 21 monsters and the spirits of birds, animals, plants and objects, and chapter 23 non-human beings that are capable of love. See also Feng Menglong’s (1986a) “Preface” to the *Shan’ge*, in *Ming Qing min’ge shidiao ji*, vol. 1, juan 1, 269–270. Moreover, other ideas elevate passions and love to a metaphysical level. In some authors the morality of *qing* could determine the ultimate status of the individual: the sincerity of *qing* could make a courtesan a paragon of virtue, far better than a wife who simply followed principle. Another cosmological justification can be found in Li Zhi, who extols the famous elopement of the widow Zhuo Wenjun with the poet Sima Xiangru, by referring to a passage of the Classic of Changes, *Yi jing* 易經 (Li Zhi 1959. 624–626).



those of Buddhist compassion, and points out its dual creative and destructive potential, and its “will to live” or “vitality” (*shengyi* 生意). Moreover, Feng’s advocacy of *qing* and its cult create a religion of *qing* moulded on Buddhism, but where paradoxically *qing* is always remembered and not forgotten: the “witness” of *qing* becomes a “passionate Tathagata of joy” (*duoqing huanxi rulai* 多情歡喜如來) for the salvation of his believers, and transforms Buddhism from a doctrine of detachment and disillusionment into a more human religion.<sup>21</sup>

## The Cult of *qing*

Just as love in the West is also the mystery of Orpheus and Eurydice, the tragedy of Romeo and Juliet, Stendhalian vanity and crystallisation, Proustian snobbishness, Flaubertian bovarysme and Karenina’s drama, so in China it is also the heroism of Sima Xiangru and Wenjun,<sup>22</sup> the immortal attachment between Han Ping and his bride<sup>23</sup> or of Panzhang and Wang Zhongxian,<sup>24</sup> the courage and determination of Yingying and Liniang,<sup>25</sup> the delicate and tormented relationship between Baoyu and Daiyu,<sup>26</sup> and other famous couples

- 
- 21 See the preface of *Qingshi* (1986b). Cf. also Hsu 1994, 166–9. It would be difficult to equate Western and Chinese ideas about the connection of love to the universe, as their substance and meanings are quite different. The Chinese form of vitalistic universalism obviously was not based on the Platonic theory of heavenly love or Galen’s theory of humours. China’s *yin-yang* dualism ventured into areas that were completely alien to the Neo-Platonic and Gnostic tension between body and soul.
- 22 Sima Xiangru 司馬相如 (ca. 179–117 BC) was a Chinese poet, writer, musician and official. His passionate love for the widow Zhuo Wenjun remained famous throughout Chinese history, both for the wretchedness and difficulties the couple had to face after escaping together and for the characters of the protagonists: their love contains a constant stream of the ingredients that make for attraction, including seduction, the woman’s transgression and elopement, humiliation and poverty. While the seduction of the young widow and the flight of the lovers had some weight in the judgement of official history, the episode immediately entered into legend; it was appropriated by the ballad singers, who have presented different versions of it over time, and it also became a favourite theme in popular theatre.
- 23 The sad story of Han Ping and his bride, He, is told in a collection of stories compiled originally in the fourth century, but whose present text comes from a Ming edition, “In Search of Prodigies”, *Soushenji* 搜神記. The king of the state of Song (period of the Warring States, 475–221 BC), having fallen in love with the beautiful He, abducts her and sends Han Ping to forced labour. In the end Han Ping commits suicide, followed by his wife, who demands to be buried beside her husband. But the king, in annoyance, buries her separately, in a tomb facing her husband’s. Miraculously, one night, a double-trunk catalpa tree grows out of both graves and within ten days or so the branches and roots from the two trunks are joined; a pair of birds come to nest among its leafage and draw the attention of the inhabitants with their sad calls. These birds are none other than the spirits of the dead couple.
- 24 The passion between Wang Zhongxian 王仲先 and Panzhang 潘章 (?–234) is an exemplary story of homoerotic love. They would be buried together after their death, and the double tree growing on their grave, whose branches are intertwined, was called “the Common Pillow Tree” (*Gongzhenshu* 公枕樹).
- 25 Two female characters of famous literary works, that represent the passionate love beyond conventions.
- 26 The main characters of the novel *Honglou meng*, famous for their unlucky love.

of tradition and literature.<sup>27</sup> Even if the self-destructive passions of Ximen in the erotic novel *JinPingMei* lack Faustian connotations, and do not display anything of the Western love/death dyad, but rather the plain retributive equation of transgression-punishment, they add a facet to the multicoloured kaleidoscope of late imperial imagery of love. The Gong'an literary circle established *qing* as the kernel of self-expression, promoted expressionistic tendencies in poetry, focused on the authenticity in expression, enhancement of the peculiarity of each writer and each time, the literary concept of “natural sensibility”, *xingling* 性靈. Occasionally, as in many erotic tales, the pursuit of pleasure is seen as a value in itself, free from any moral judgement, but limited by the sense of *horror* for any excess or dissoluteness.

By syncretically combining passionate love with the virtue of humaneness (*ren*) and with the inborn but subjective conscience (*liangzhi* 良知) of Wang Yangming, the writers of this trend defined the term *qing* as a generally positive and multi-faceted feeling that expressed a wide range of emotions: empathy for one's fellow man, Confucian virtues – starting with filial piety and respect for one's superiors – a profound sense of universal humaneness linking an individual to all the other beings and things in the cosmos, like the ribbon used to string coins together; that vital energy which represents the driving force behind the continuous production and reproduction of the universe; although corresponding to a personal feeling it was supposed to be more powerful than vices and virtues.<sup>28</sup> Sublimation and idealisation of love through a number of forms of expression, the crisis of traditional gender roles, a heightening of the aesthetic and narcissistic pleasure in contemplating one's own “flow of consciousness” are all symptoms of an evolution in the Chinese imaginary world during the last two dynasties, before the advent of Western influence.

We talk of a “cult of *qing*”, because it is possible to notice a new trend in philosophical and moral fields, as well as in literary production, committed to re-discussing the role and evaluation of emotions and desire, the general and the particular. Sources demonstrate the emergence of a great debate on the role of *qing*, corresponding to a real “cult”, based on the common evaluation and appreciation of *qing*, even if the literati who promoted it did not necessarily share a common programme and did not intend to create a uniform movement with a unique ideology; rather different intellectual currents met in it. The “cult of *qing*” is a key concept for understanding the ideological impact on late imperial society and its

---

27 On the transformation of literary characters into real personages, like Liniang, see Widmer 1989, 44–5. On the fiction–reality relationship, cf. Grass 1985, 14–5.

28 See the introduction probably written by Feng Menglong (1986b), under the pseudonym of “Long Ziyou of Wu”, to the “History of Love”, and the comments at the end of each chapter (especially juan 20, 694). A similar attitude is also found in other works of the same period such as Tang Xianzu's “Peony Pavilion”, the twenty-fifth story of Ling Mengchu's “Slapping the Table in Amazement” (*Pai'an jingqi*) or the fourteenth in the “Second Collection of the Western Lake” (*Xihu erji*). “For all the men on earth,” says a monk in the drama *Huazhongren*, “if love is sincere, separated lovers can be reunited and the dead brought back to life.” (cit. by Yao 1982, 47). To be recalled is also the mentioned ideological contribution of Wang Yangming and the Taizhou School, with the transition from the transcendence of the principle to innate moral conscience.

polyphonic contributions. The common re-evaluation of *qing*, however, concurs in establishing its transcending social and gender limits and reconceptualising the human person basically in terms of *qing*.

Precisely because of its innovative implications, in its negotiation with the dominant Confucian morality, the cult of love nevertheless required moral justification in order to be accepted as a value per se. Identification of emotions with principles in order to find a moral support for the sentiments was the basis of the positive attitude to emotions which was complementary to the restrictive cautiousness toward them. The above was the first device to rehabilitate emotions and desires, using the oxymoron from the Daoist reformistic moral arguments: “following the principle through the emotions” (以情從理).<sup>29</sup>

The second device consisted of the vitalistic arguments of the dynamic yin-yang interaction and a certain sacralisation of love. These devices paved the way to elevate *qing* to a metaphysical level close to the concepts of heaven, destiny, tracing back to the cosmic process of generation and regeneration, and overcoming the limits of life and death.

The search for *genuine* emotions was another argument in favour of passions, as it leads back to the concept of a close relation between human nature (*xing*) and affections (*qing*). The metaphors “love is like an itch” and “love is like food” that can summarise the “minimalist” moral of the song collection “Mountain Songs” represent the basic understanding of sexual love in its essence. Feng Menglong, in his introduction to this collection, contrasts the sincerity of these feelings and the “false medicine” of the hypocrisy of social conventions; while Tang Xianzu goes as far as to state that the moralism of principles does not allow human feelings to be understood.<sup>30</sup> Recognised as genuine was the concrete nature of love in its double functions of fulfilment of human appetites and social needs (procreation and family offspring). It is clear otherwise that love could not be perceived only as concrete need. From the ideological point of view Confucian rules aimed at channelling libido energies within the family organisation, and Daoism – where the idea of *qing* is not anchored to social, economic and procreative concerns, and sex is not bound to reproductive functions – aimed to cultivate them according to hygienic body care. But passion’s inebriating effects requested more explanations beyond the minimalist love representation. Some writers elaborated an idea of love similar to a supernatural experience, a yearning for infinite and eternal values, in a kind of religious dimension. They, however, had to cope

---

29 This expression was used in the debate between Wang Bi 王弼 (ca. 226–249) and He Yan 何晏 (190–249), and may be interpreted in various ways (eliminate emotions by principle, following principle through emotions, reacting emotionally without being ensnared by things). He Yan, following literally *Zhuangzi* (third century B.C.), excluded that the sage could have *qing* (聖人無喜怒哀樂), while Wang Bi maintained the opposite thesis: the sage was superior to other men in terms of wisdom, but he was like other men in his having the five emotions (同於人者五情也): “The emotions of the sage are such that though he reacts to things, he is not ensnared by them (然則聖人之情，應物而無累於物者也). It is a great error, consequently, to say that because he is not ensnared by things, he therefore has no (emotional) reactions to them.” (cf. Fung, Bodde tr., 1953, vol. 2, 188–9).

30 Feng Menglong (1986a), Foreword to the *Shan’ge* (Songs of the Mountains), juan I, 269. Tang 1962 2, juan 33, 1093. See also Lowry 2005. 142–328.

with human limits in time and space, inner and external changes, besides the pressing of social needs. The literary genre called “talent and beauty” responded to both orthodox and sentimental expectations, seconding the cult of *qing* in giving space to the elevation of love and passions, and offering arguments in favour of its value, morality and merits. But several works seem to present an inconsistency, which varies from author to author, a sign of a malaise that consciously or unconsciously afflicts the writers of this period, which offers plural spaces for creation, negotiation and subversion.<sup>31</sup> The problem may be schematically outlined as a conflict of values, the late Ming conflict between traditional morality – which very few dare deny formally – and the value of love and passions per se, or the emphasis on one of various aspects of love.

A composite sentiment like love is manifested with several experiences which have analogies even in different societies; yet the basic drive aimed at reproduction acquires a multiplicity of complex aspects elaborated by each culture and evolving with the times: the rite of courting, seduction rhetoric, lovesickness, jealousy, institutionalisation of sexual relation, etc. What is different is the way such phenomena are elaborated, represented and lived. Philosophical and moral elaborations as well as literary expressions offer a rich material for reading social and subjective interpretation and justification of these elements.

Generally speaking, *qing* was not extolled above all values, or countered to society and conventions. Haiyan Lee stresses that

the ‘cult of *qing*’ does not effect an epistemic break with Neo-Confucian orthodoxy. For all its effort to legitimise the affective and the individual, it is still committed to patrilineal continuity, ritual propriety, and the social order. The supremacy of ritual is rarely questioned and sexuality is rarely affirmed on the ground of the leisure principle. All social relationships are made transubstantial through *qing* and yet are still regulated by *li*. With few exceptions, the texts produced in the movement do not antagonise the cardinal relationships, or pit the horizontal axis of sentimentality against the vertical axis. Conjuality seldom stands alone without filiality and sexual passions are laudable only if they also validate, if not actually strengthen, the parent-child bond. (Lee 2007, 38; 2001, 291–327)

We cannot but agree regarding the important role played by the “Confucian sentimentality” that corresponds to “good sentiments” or “moral sentiments”. Virtues were somehow subordinated to sentimental logic, but most of the contents remained anchored to the same

---

31 For instance, Mei Chun (2009, 26) raises this issue concerning a story of Feng Menglong: “The Pearl Shirt’ poses an interpretive conundrum: not only is the narrator inconsistent, but readers are at a loss about Feng Menglong’s own standpoint, since he seems to espouse both moralistic retribution and genuine feeling (*qing* 情). [...] The interpretive inconsistency, or discrepancy, for modern readers arises from Feng’s conscious usage of two narrative spaces, the prose space and the verse space, to juxtapose a highly moralizing narratorial voice with a countervoice influenced by Wang Yangming’s ‘School of the Mind’. [...] Through the hybrid voices of verse and prose, ‘The Pearl Shirt’ produces multiple layers of storytelling, moralizing, counter-moralizing, and reading.”

Confucian virtues, i.e. social values that were validated as far as they were performed with genuine feelings. Few Chinese writers stressed the conflicts between social conformity and individual desires. It seems that in China social expectations for individual behaviour have been ultimately more effective than the sexual rules of the Christian Church in Europe, notwithstanding their menace of eternal damnation (Ropp 2006, 226 and Santangelo 1991). On the other side I would be cautious in underrating the significance of new values and perspectives inherent to the cult of *qing* and its developments in the sixteenth, seventeenth and eighteenth centuries. They did not become a dominant vision of the individual and of love as being a direct opponent of tradition, and did not aim at highlighting individual autonomy as against social conventions. Nevertheless, by stressing the central role of emotions, the value of love and beauty, and extolling the universality of desires, the cult of *qing* allowed an expansion of the self perception that did not reject the five cardinal relations, but considerably reformed their hierarchic order.

## Chinese Culture of Love in a Comparative Perspective

If compared with European civilisation, categories and elaborations used in Chinese culture are obviously different. Chinese poems and narrative, for instance, lack any conceptions equivalent to Eros and Aphrodite, and are in contrast with the prevalent anthropomorphism in Europe, from Hellenistic poetry to Romanticism. The allegories of the mirror and the dream are present in Western and Chinese civilisations with a multiplicity of meanings, and may be applied to passions, virtues and vices. Indeed, the exaltation of the emotions in the late Ming period makes extensive use of these motifs to assert the transcendence of the passions and enrich the world of desires and illusions. Nevertheless, while in Europe the fantastic and metaphorical element was extolled as the supreme stage of the phenomenon of love, in Chinese culture the same fantastic element was considered an allegory for the imaginary, fleeting, unreal nature of passion and desire.

The cultural, political, ideological and religious context of the growth and development of the cult of *qing* is completely different from the context from which the Renaissance and later European Romanticism were born and developed. If we take into account the conventional periodisation and the refined arguments offered by the research done in recent decades, the change of perspective in perceiving emotions after the European Middle Age is huge. The Medieval sublimation of love was based on the religious-metaphysical belief centred on a transcendent God that made man, and thus shared a body with human beings, with the same physical and moral feelings – from hunger to pain, from indignation to affection, agony and death and resurrection. In this Christocentric vision, love was ideally legitimated and nourished by the Love of God, and passions were accepted and sanctified by the Passion of Christ; thus all emotions were tendentially transfigured by faith (Boquet

and Nagy 2018). The passage from this system to modernity involves the separation of natural sciences from theological–religious beliefs; individuals and communities became independent from the normative concept of love commanded by a transcendent Creator. The Chinese did not share the fear regarding the body and carnal sin of Christianity, but still they developed a fear of inner and outer unbalance, hygienic and social rules. Sexual freedom was not a carnal sin, rather a threat to social morality and customs. If we restrict our analysis to the possible *analogy* of love in the two contexts, we can say that analogy consists of “identity with nature, emphasis on the originality and spontaneity, plea for creative sensibility, and the concept of the urgency and overflow of feeling” (Hung 1974, 2, footnote 5); nevertheless, we are well aware that the semantic contents of the notions “nature”, “spontaneity”, “creative sensibility” and “overflow of feeling” do not correspond in the two cultures (Chaves 1985, 123–50). Moreover, the role of religion, the relationship between morality and aesthetics, the hierarchy of values do not match in the two cultures, and thus these conditions influenced the way of perception and evaluation of emotions and its evolution. In this sense, we can recognise the birth of a real cult of *qing*, but on the basis of ideological and cognitive elements different from those behind the movements developed in Europe.

In the last century Michael Balint (1896–1970) proposed two fundamental models for classifying the different forms of civilisation based on the attitude to eros: 1) societies in which passionate love is extolled, its object idealised, courtship is important and love poetry and songs have particularly developed, together with the language of seduction; 2) societies that favour pure genital sexuality and neglect the mediation of idealisation and tenderness, such as love poetry and songs, and courtship (Balint 1991, 118). It would be difficult to classify Chinese culture in the first or in the second model, as in China extreme pragmatism is combined with the need to discover a supernatural aura around it. For instance, Zhou Ruchang argues that “Chinese culture is the culture of sentiments”, *qingwenhua* 情文化 (quoted from Lee 1997, 110). Among various representations of love in literary and philosophical sources we can appreciate – borrowing Zhu Guangqian’s (1897–1986) expression – some of the infinite ways of “covering a very trivial thing with an extremely refined façade, that is something “outside shining gold and jade, inside cotton waste” “它是替一件極塵俗的事情，掛上一個極高尚的幌子，「金玉其外，敗絮其中」”。 (Zhu Guangqian 1987, 109–111). What is called “trivial” in fact is the basic phenomenon of evolving life in the universe, and one of the essential motivations of individual experiences, elation and despair. This sense of decency, notwithstanding the development of the cult of *qing*, is probably due to the tradition of making use of a mediated and symbolic language, rather than to the legacy of the dominant cautiousness concerning any sexual excess. After all, the long-lasting naturalistic tradition from Gaozi 告子 (ca. 420–350 BC) to He Xinyin 何心隱 (1517–1579) could not have been swept away by a few decades of revolutionary fervour.<sup>32</sup>

32 Well known is Gaozi’s statement “Appetite for food and desire for sex are part of human nature” (食色性也), from which we can infer that sexual drive was considered a human natural need with hunger and thirst (*Mencius*, Gaozi. 1). Thus we may understand the future conflict and tension between erotic attraction as lustful tendency, *yin* 淫/*yu* 欲, and the genuine and true feelings, *zhenqing* 真情;



In conclusion, the so-called cult of *qing* is a phenomenon developed in late imperial China between the sixteenth and eighteenth centuries. It was propagated both among thinkers and literary writers, and aimed at re-evaluating passions and desires that Neo-Confucian orthodoxy contrasted with moral principles. Reflecting the great changes and developments in the social and economic fields, the School of Mind undermined the principle of authority of orthodox Neo-Confucianism. The reduction and transformation of the role of principles allowed more space for the legitimacy of emotions and feelings.

The cultural and social context of the growth and development of the cult of *qing* is completely different from the context from which the Renaissance, Enlightenment and later European Romanticism were born and developed. It developed from orthodox and heterodox ideas from Buddhism, Daoism, Neo-Daoism, Confucianism and popular beliefs. A debate about the “genuineness” and “naturalness” of *qing* accompanied the new trend. The search for *genuine* emotions was the first argument in favour of passions, as it led back to the ancient concept of the close relation between human nature (*xing*) and affections (*qing*). This meant a re-evaluation of passions justified by their inner moral value, and identified with principles. Moreover, a vitalistic argument linked emotions to the dynamic of yin-yang interaction and a certain sacralisation of love.

The Neo-Confucian polarities principle-passions and excess-moderation followed the rhetoric of the “egoism-universalism” dyad (*si-gong* 私公), in which the second term does not indicate only what is public and general, but also what is objective, in harmony with the world, and not bent to the particular. In the late Ming dynasty, however, these ways of understanding them were challenged. Once argued that desires definitely derive from heaven (innate and authentic), what was usually branded as selfish could be considered from a different perspective: in the ambit of Neo-Confucianism the concept of “real” human nature emerged together or in contrast with the traditional Confucian moral nature, and this allowed a new morality on the basis of the extension of individual inclinations to others by reciprocity and tolerance. According to some thinkers of the so-called Taizhou School, the self consists of natural needs, which are physical as well as self-centred inclinations. “Natural self” corresponds to genuine human nature, neither good nor bad, with the basic individual’s desires: appetite for food, attraction for beauty, shelter from the cold, desire for elegance, quest for comfort, the fulfilment of one’s ambitions. Individual autonomy, self-respect, self-love, self-interest and self-preservation in fact were no more egoistic, as they meant the fulfilment of one’s needs as well as those of others, through a process of universalisation of desires.

The rise of a cult of passions in its various forms is a symptom of an evolution in the Chinese imaginary world during the last two dynasties, before the advent of Western influence. We notice a new trend in philosophical and moral fields from the orthodox contrast

---

both are passions and desires, but the former have a negative connotation, the latter a positive one. The famous sentence of He Xinyin states that desires, the enjoyment of food, the pleasures of love and the joy of music are part of human nature 性而味, 性而色, 性而聲 (*Cuantong ji* 爨桐集, in *He Xinyin ji* 何心隱集, cit. in Rong 1941, 229).



between principles and desires to a moderate re-evaluation of passions and desires, as well as in literary production, committed to re-discussing and renegotiating the role and evaluation of emotions and desire, the collective and the peculiar. Provided *qing* was neither extolled above all values or countered to society and conventions, nor contrasted with Confucian tradition, in the new trend social values were validated as far as they involved “genuine feelings”. Positing the centrality of emotions, the power of love and beauty, and celebrating the universality of feelings and desires, the cult of *qing* opened the way to alternative anthropological visions that did not antagonise the cardinal relations, but led towards a new way to live them.

## Bibliography

For references to ancient Chinese texts and Classics, if no mention is given for specific editions, see *Chinese Text Project*, online <http://ctext.org/>.

### Primary Sources

- Aristotle. 2014. *The Nicomachean Ethics*. Edited by Christopher Shields. London – New York: Routledge. <https://oll.libertyfund.org/titles/aristotle-the-nicomachean-ethics>
- Cao, Xueqin 曹雪芹 and Gao E 高鶚. 1998. *Hongloumeng* 紅樓夢 (*Chengjia* 程甲本 edition printed in 1791). Beijing: Zhonghua shuju.
- Feng, Menglong. 1986a. *Shan'ge* 山歌, by [after 1618], repr. in *Ming Qing min'geshitiao ji* 明清民歌詩調集, Shanghai: Shanghai guji chubanshe, neibu.
- Feng, Menglong. 1986b. *Qingshi leilue* 情史類略, [abridged: *Qingshi* (History of Love)]. Shenyang: Chunfeng wenyi chubanshe.
- Hawkes, David and Minford, John (trans.). 1979–1986. *Cao Xueqin, The Story of the Stone, also Known as The Dream of the Red Chamber*, 5 vols. Harmondsworth: Penguin Books.
- He, Xinyin 何心隱. 1981. *He Xinyin ji* 何心隱集 (Works of He Xinyin), Rong Zhaozu 容肇祖, ed., repr. Beijing: Zhonghua shuju.
- Li, Zhi 李贄. 1959. *Cangshu* 藏書 [A book to keep]. Beijing: Zhonghua shuju.
- Li, Zhi 李贄. 1975. *Fenshu* 焚書 [A book to burn]. Beijing: Zhonghua shuju.
- Ling, Mengchu 凌濛初. 1982. *Pai'an jingqi* 拍案驚奇 [*Slapping the Table in Amazement*]. Shanghai: Shanghai guji chubanshe.
- Luo, Qinshun 羅欽順. 1999. *Kunzhiji* 困知記 [Knowledge Painfully Acquired]. In *Siku quanshu buzheng* 四庫全書補正, Zibu 子部, Rujia lei 儒家類, *juan xia* 卷下, repr. Taipei: Taiwan Shangwu.
- Ling, Mengchu 凌濛初. 1982. *Pai'an jingqi* 拍案驚奇 [Slapping the Table in Amazement]. Shanghai: Shanghai guji chubanshe.
- Luo, Qinshun 羅欽順. 1999. *Kunzhiji* 困知記 [Knowledge Painfully Acquired]. In *Siku quanshu buzheng* 四庫全書補正, Zibu 子部, Rujia lei 儒家類, *juan xia* 卷下, repr. Taipei: Taiwan Shangwu.
- Mencius*. Online edition in *Chinese Text Project*. <https://ctext.org>

- Plato. *Laws*. Translated by Benjamin Jowett. <https://ebooks.adelaide.edu.au/p/plato/p711/complete.html>
- Siku Quanshu* (Complete Library of the Four Treasuries, SKQS), Jingyin Wenyuange Siku Quanshu 景印文淵閣四庫全書, Taipei: Taiwan Shangwu yinshuguan, available in electronic edition at the *Wenyuange Siku Quanshu* 文淵閣四庫全書電子版.
- Tang, Xianzu 湯顯祖. 1962. *Tang Xianzu Ji* 湯顯祖集 [Tang Xianzu's Collected Works]. 4 vols. Beijing: Zhonghua shuju.
- Wang, Gen 王艮. 1975. *Wang Xinzhai quanji* 王心齋全集 [Wang Gen's Complete Works], repr. Taipei: Guangwen shuju 廣文書局. <https://ctext.org/wiki.pl?if=en&res=124490>
- Wang, Yangming 王陽明. *Yangming quanshu* 王陽明全書 [Wang Yangming's Complete Works]. In *Siku Quanshu* electronic ed.
- Xenophon. *The Memorabilia. Recollections of Socrates*. Translated by H. G. Dakyns. <http://www.gutenberg.org/files/1177/1177-h/1177-h.htm>
- Yuan, Haowen 元好問, *Yanqiu ci* 雁丘詞, in *Yishan yuefu* 遺山樂府 [Yuan Haowen's Ballads]. In *Siku quanshu* electronic ed., Jibu 集部, Cidian lei 詞曲類, Cixuan zhishu 詞選之屬, *Yu xuan lidai shiyu* 御選歷代詩餘, *juan* 119. Shanghai Renmin chubanshe, n.d.
- Zhu, Xi 朱熹. 1983. *Zhuzi yulei* 朱子語類 [Classified Dialogues of Master Zhu]. Beijing: Zhonghua shuju. <http://ctext.org/zhuzi-yulei/10>

## Secondary Sources

- Balint, Michael. 1991. *L'amore primario. Gli inesplorati confini tra biologia e psicoanalisi* [Primary Love and Psycho-analytic Technique, 1952]. Milano: Raffaello Cortina.
- Boquet, Damien and Nagy Piroska 2018. *Medieval Sensibilities. A History of Emotions in the Middle Ages*. Cambridge: Polity Press.
- Chaves, Jonathan, 1985. "The Expression of Self in the Kung-an School: Non-Romantic Individualism." In *Expressions of Self in Chinese Literature*, edited by R. Hegel and R. Hessney. New York.
- Cheng, Yu-yin. 2013. "Changing Cosmology, Changing Perspectives on History and Politics: Christianity and Yang Tingyun's 楊庭筠 (1562–1627) Reflections on China." *Journal of World History* 24(3): 499–537.
- de Bary, William Theodore. 1970. "Individualism and Humanitarianism in Late Ming Thought." In *Self and Society in Ming Thought*, edited by W. T. de Bary. New York: Columbia University Press.
- Epstein, Maram. 2001. *Competing Discourses: Orthodoxy, Authenticity, and Engendered Meanings in Late Imperial Chinese Fiction*. Cambridge: Harvard University Press.
- Fung Yu-lan [Feng Youlan]. 1952–1953. *A History of Chinese Philosophy*. Edited and translated by D. Bodde, 2 vols. Princeton: Princeton University Press.
- Grass, Gunter. 1985. "Fictions Are Lies That Tell the Truth – Salman Rushdie and Gunter Grass in Conversation." *The Listener* 27 June.
- Hanan, Patrick. 1981. *The Chinese Vernacular Story*. Cambridge, MA: Harvard University Press.

- Hansen, Chad. 1972. "Freedom and Moral Responsibility in Confucian Ethics." *Philosophy East and West* 22: 169–86.
- Heller, Ágnes. 1985. *Il potere della vergogna* [The Power of Shame: A Rational Perspective]. Rome: Editori Riuniti.
- Honda, Wataru 本田濟. 1983. "Shujō no setsu – Shinchōjin no baai" 主情の説 – 清朝人の場合, in Kanaya Osamu 金谷治, ed., *Chūgoku ni okeru ningensei no tankyū* 中国における人間性の探究. 643–63. Tokyo: Shōbunsha 創文社.
- Hsia, C. T. 1970. "Time and the Human Condition in the Plays of T'ang Hsien-tsu." In *Self and Society in Ming Thought*, edited by W. T. de Bary. New York: Columbia University Press.
- Hsieh, Daniel. 2007. *Love and Women in Early Chinese Fiction*. Hong Kong: Chinese University Press.
- Hsu, Pi-ching. 1994. *Celebrating the emotional self: Feng Meng-lung and the late Ming ethics and aesthetics*. P. D. Dissertation, University of Minnesota.
- Huang, Weizong, Martin. 1998. "Sentiments of Desire: Thoughts on the Cult of Qing in Ming-Qing Literature," *CLEAR* 20: 153–184.
- Huang, Weizong Martin. 2001. *Desire and Fictional Narrative in Late Imperial China*. Cambridge Mass.: Harvard University Press.
- Huang, Zhouyue 黄卓越. 2012. *Sixiangshi yu pipinxuelun wenji* 思想史与批评学论文集. Beijing: Beijing Language and Culture University Press.
- Huang, Zongxi 黄宗羲. 1984. *Mingru xue'an* 明儒學案 (*Treatise on the Ming scholars*). Taipei: Shijie shuju.
- Hung, Ming-shui 1974. *Yüan Hung-tao and the Late Ming Literary and Intellectual Movement*. Ph. D. Dissertation, University of Wisconsin.
- Jiang, Guozhu 姜國柱 and Zhu Kuiju 朱葵菊. 1989. *Zhongguo lishi shang de renxing lun* 中國歷史上的人性論. Beijing: Zhongguo shehuikexue chubanshe.
- Jongman-Sereno, Katrina P. and Leary, Mark R. 2018. "Self-judgments of authenticity." *Self and Identity*: 1–32.
- Lee, Haiyan. 1997. "Love or Lust? The Sentimental Self in Honglou meng." *CLEAR* 19: 85–111.
- Lee, Haiyan. 2001. "All the Feelings That Are Fit to Print. The Community of Sentiment and the Literary Public Sphere in China, 1900–1918." *Modern China* 27(3): 291–327.
- Lee, Haiyan. 2007. *Revolution of the Heart: A Genealogy of Love in China, 1900–1950*. Stanford: Stanford University Press.
- Li, Wai-ye, 1993. *Enchantment and Disenchantment. Love and Illusion in Chinese Literature*. Princeton: Princeton UP.
- Li, Wai-ye. 1995. "The Rhetoric of Spontaneity in Late-Ming Literature." *Ming Studies* 35: 32–52.
- Lowry, Kathryn. 2005. *Tapestry of Popular Songs in 16th and 17th Century China. Reading, Imitation, and Desire*. Leiden: Brill.

- Lu, Ji 陸楫. 1937. *Jianjiatang Zazhu Zhaichao* 蒹葭堂雜著摘抄 [Excerpts and Miscellanea from the Hall of Reeds and Rushes]. In *Congshu Jicheng Chubian* [First Edition of the Collected Series of Books], edited by Wang Yunwu, First Part, Item 9.
- Mei, Chun. 2009. "Garlic and Vinegar. The Narrative Significance of Verse in 'The Pearl Shirt Reencountered'." *Chinese Literature: Essays, Articles, Reviews* 31: 23–43.
- Oki, Yasushi and Paolo Santangelo. 2011. *Shan'ge, the 'Mountain Songs'. Love Songs in Ming China*. Leiden: Brill.
- Rong, Jiaozu 容肇祖. 1941. *Mingdai sixiang shi* 明代思想史 [History of Ming Thought]. Shanghai: Kaiming shudian.
- Ropp, Paul. 2006. "The Price of Passion in Three Tragic Heroines of the Mid-Qing: Shuangqing, Lin Daiyu, and Chen Yun." In *From Skin to Heart. Perceptions of Bodily Sensations and Emotions in Traditional Chinese Culture*, edited by Santangelo Paolo, in cooperation with Ulrike Middendorf, 203–28. Wiesbaden: Harrassowitz.
- Ruan, Jingdong 阮景东, 2014. *Diguo de qimeng: Mingchao de lishi liebian* 帝国的启蒙: 明朝的历史裂变 [Imperial Enlightenment: Ming Historical Revolution], Hangzhou: Hangzhou Chuangce keji youxiangongsi 杭州创策科技有限公司. <http://vip.book.sina.com.cn/weibobook/vipc.php?bid=5352429&cid=10282295>
- Santangelo, Paolo. 1991. *Il "peccato" in Cina. Bene e male nel neoconfucianesimo dalla metà del XIV alla metà del XIX secolo*. Bari: Laterza.
- Santangelo, Paolo [Shi Hualuo 史華羅]. 2000a. "Zhongguo yu Ouzhou 'aiqing' gainianhua de zongjiao yinxiang" 中國與歐洲 '愛情' 概念化的宗教影響 [Influence of Christianity and Chinese Tradition on the Conceptualization of Love]. *Journal for the Study of Christian Culture* 基督教文化學刊 4: 41–71.
- Santangelo, Paolo. 2000b. "The Cult of Love in Some Texts of Ming and Qing Literature." *East and West* 50: 439–499.
- Santangelo, Paolo. 2006. "Reconsidering the 'Cult of qing' in Late Imperial China: A 'Romantic Movement' or a Conveyer of Social Values?" *MingQing yanjiu*: 133–163.
- Santangelo, Paolo. 2016. *L'uomo fra cosmo e società. Il neoconfucianesimo e un millennio di storia cinese*. Milano: Mimesis.
- Santangelo, Paolo. 2017. "The Literati's Polyphonic Answers to Social Changes in Late Imperial China." *Frontiers of History in China* 12(3): 357–432.
- Santangelo, Paolo and Boros Gábor. 2019. *The Culture of Love in China and Europe*. In press by Brill, ESMA series.
- Sedikides, Costantine, Slabu, Letitia Gabriela, Lenton, Alison, and Sander, Thomas 2017. "State authenticity". *Current Directions in Psychological Science* 26(6): 521–5. [https://www.researchgate.net/publication/316862597\\_State\\_Authenticity](https://www.researchgate.net/publication/316862597_State_Authenticity)
- Sun Chang, Kang-i. 1990. *The Late Ming Poet Ch'en Tzu-lung: Crises in Love and Loyalism*. New Haven: Yale University Press.
- Wang, Gang 王崗. 1999. *Langman qinggan yu zongjiao jingshen: Wanming wenxue yu wenhua sichao* 浪漫情感與宗教精神: 晚明文學與文化思潮 [The Romantic Sentiment and the Religious Spirit: The Late Ming Literature and the Intellectual Currents]. Hong Kong: Tiandi Tushu Cosmos Books.

- Wang, Richard. 1994. "The Cult of *Qing*: Romanticism in the Late Ming Period and in the Novel *Jiao Hong Ji*." *Ming Studies* 33.
- Widmer, Ellen. 1989. "The Epistolary World of Female Talent in Seventeenth-Century China." *Late Imperial China* 10(2).
- Xu, Fuming 徐扶明 (ed.). 1987. *Mudanting yanjiu ziliao kaoshi* 牡丹亭研究資料考釋. Shanghai: Shanghai gujichubanshe.
- Xu, Hong 徐泓, 1986. "Mingmo shehui fengqi de bianqian-yi Jiang-Zhe diqu wei li" 明末社會風氣的變遷以江浙地區為例 [The Changes in Late Ming Social Customs—A Case Study of the Jiangzhe Area]. *Dongya wenhua* 東亞文化 (East Asian Culture) 24: 83–110. Taipei.
- Yao, Christina Shu-hwa. 1982. *Cai-zi jia-ren: Love Drama During the Yuan, Ming and Qing Periods*. Ph. D. Dissertation, Stanford University.
- Zhou Daxing 周大興. 2000. "Wang Bi 'Xing qi Xing' de Renxing Yuanjin Lun" 王弼「性其情」的人性遠近論 [Wang Bi's Discourse on the Proximity of Human Nature 'Naturalisation of Emotions']. *Zhongguo Wenzhe Yanjiu Jikan* 中國文哲研究集刊 16: 339–74.
- Zhou, Ruchang 周汝昌. 1993. "Honglou meng yu 'qingwenhua'." 紅樓夢與情文化, *Honglou meng xuekan* 紅樓夢學刊 1.
- Zhu, Guangqian 朱光潛, 1987. *Tan xiuyang* 談修養. In *Zhu Guangqian quanji* 朱光潛全集, 4, Hefei: Anhui jiaoyu chubanshe.



# List of contributors

**Gábor Boros** is professor of Philosophy (ELTE), and of General Humanities (Károli Gáspár University of the Reformed Church in Hungary). He is the president of the Hungarian Philosophical Association. He has published extensively on Early Modern Philosophy (Descartes, Spinoza, Leibniz), the philosophy of emotions in general and of love in particular. He is director of the series *Rezonőr* established to publish translations of the most significant works of and about francophone philosophy.

**Ákos Forczek** is a PhD student in philosophy at Eötvös Loránd University Budapest (ELTE). His main research area is Kant's theory of experience, paying special attention to the function of imagination. He is writing his dissertation under the supervision of Nóra Szegedi and Gábor Boros.

**Heikki Haara** is a senior lecturer at the University of Helsinki. He gained his PhD in Helsinki and has been a visiting researcher at the universities of California, Berkeley and Oxford. He has worked on early modern moral and political thought, particularly on natural law and theories of sociability and social emotions. His publications include *Pufendorf's Theory of Sociability: Passions, Habits and Social Order* (Springer, 2018) and several articles in journals including *Journal of the History of Ideas*, *Political Theory* and *Journal of Scottish Philosophy*.

**Przemysław Gut** is an associate professor of philosophy at the John Paul II Catholic University of Lublin. He has published in the field of the history of early modern philosophy. He is the author of *Spinoza o naturze ludzkiej* [*Spinoza on Human Nature*] (KUL Publishing House, 2012), numerous articles in early modern philosophy and a contributor to *The Radical Enlightenment*, *Diametros* 40 (2014). Currently, he is writing a book on Descartes.



**Mariangela Priarolo** is a lecturer in the history of philosophy at the University “Ca’ Foscari”-Venice (Italy). Her research has focused on the relationship between metaphysics, theology and epistemology, especially in Malebranche and Leibniz. Her current interests are the connection between metaphysics and science in the seventeenth century, and the place of rationalism in the history of toleration. She is the author of *Il determinismo. Storia di un’idea* (Determinism. The History of an Idea, Rome 2011) and *Malebranche* (Rome, 2019).

**Paolo Santangelo**, Professor Emeritus of “Sapienza” University of Rome (East Asian History), is editor of the yearly publication *Ming Qing Studies*, is Director of the series *Emotions and States of Mind in East Asia*, of Brill publishers, and leads an international research project on the textual analysis of literary and non-literary sources in Chinese culture. He has published several volumes and articles on Chinese anthropological history and on the culture of late imperial China.

**Wilhelm Schmidt-Biggemann** is Professor Emeritus of History of Philosophy and the Humanities at the Freie Universität Berlin, and since 2016 he has been a distinguished Friedrich Senior Professor at the Freie Universität. He was the co-founder and speaker (2003–2006) of the interdisciplinary Forum Middle Ages – Renaissance – Early Modernity at the Freie Universität. He was speaker of the DFG Research Group “Topics and Tradition” (Topik und Tradition) from 2005 to 2010. From 2008 to 2016 he was a member and since 2012 also the speaker of the Review Board ‘philosophy’ of the German Research Association (DFG). He has taken up numerous fellowships and guest professorships in Prague, Princeton, Copenhagen, Cambridge, Philadelphia and Tel Aviv. He is a recipient of the Comenius Medal and the Golden Medal of Karls University of Prague. In 2013 he was awarded the Johann Georg Hamann-Forschungspreis of the University of Münster. He has published widely on the philosophy of religion, the philosophy of history, the history of philosophy and philology, especially of the Early Modern period, political theology and the history of the Christian Kabbalah.

**József Simon** is Associate Professor at the Department of Philosophy, University of Szeged. His research topics include scholastic, renaissance and early modern metaphysics, epistemology, political philosophy and philosophy of religions, as well as the history of philosophy in early modern Hungary.

**Judit Szalai** teaches at Eötvös Loránd University, Budapest. Her main interests are seventeenth-century thought, the philosophy of mind and philosophical psychology. Her recent publications are on the emotions, rationality/irrationality, and the non-conscious in historical and contemporary contexts.

**Szilárd Tattay** is associate professor at the Department of Legal Philosophy of Pázmány Péter Catholic University Faculty of Law and Political Sciences in Budapest. From 2013 to 2018 he was a research fellow at the Institute for Legal Studies of the Centre for Social Sciences of the Hungarian Academy of Sciences. He is a founding member of Somló Bódog Society. He has been a visiting research fellow at Université Saint-Louis, Brussels, at Katholieke Universiteit Leuven and at the University of Notre Dame, Indiana. His research interests cover the history of legal and political ideas, legal philosophy, theory of the state and political philosophy.

**Peter West** is PhD candidate at Trinity College Dublin. His research is motivated by the philosophical question of what it means for one thing to represent another and the historical question of what role debates about representation, and associated issues like intentionality and language, played in the progress of Early Modern thought. His paper 'Anti-representationalism in Berkeley and Sergeant' won the 2019 International Berkeley Essay Prize, and his most recent publication 'Reid and Berkeley on Scepticism, Representationalism, and Ideas' appeared in the *Journal of Scottish Philosophy* in 2019.

**Charles T. Wolfe** is an Assistant Professor in the Dipartimento di Filosofia e Beni Culturali, Università Ca'Foscari, Venice (and project ERC EarlyModernCosmology, GA 725883). He works primarily in history and philosophy of the early modern life sciences, with a particular interest in materialism and vitalism. He is the author of *Materialism: A Historico-Philosophical Introduction* (2016) and *La philosophie de la biologie: une histoire du vitalisme* (2019), and has edited volumes including *Monsters and Philosophy* (2005), *The Body as Object and Instrument of Knowledge* (2010, with O. Gal), *Vitalism and the scientific image in post-Enlightenment life-science* (2013, with S. Normandin), *Brain Theory* (2014), *Physique de l'esprit* (with J. C. Dupont, 2018), and *Philosophy of Biology before Biology* (with C. Bognon-Küss, 2019). He also co-edits the Springer series in History, Philosophy and Theory of the Life Sciences. Papers available at <http://ugent.academia.edu/CharlesWolfe>.

**Bartosz Żukowski** is Assistant Professor at the Department of the History of Philosophy, Institute of Philosophy, University of Łódź, Poland. His research focuses on early modern idealism, philosophy of mind, philosophy of language and structural semantics. He is the author of *Esse est percipi? George Berkeley's Metaphysics of Idea* (2012, in Polish) and *Unnoticed Revolution. Richard Burthogge's Constructivist Idealism* (2019, in Polish).



With the publication of the papers of the First Budapest Seminar in Early Modern Philosophy in 2017, we accomplished overcoming such unfruitful divisions as the time-honoured interpretive distinction between “rationalists” and “empiricists”, liberating our perspectives from the rigid prejudices of simplifying handbooks. We also prepared the frame for further in-depth investigations in other areas of Early Modern thought, such as are presented in our volume. It comprises papers based on the contributions to the Second and Third Budapest Seminar in Early Modern Philosophy, held on 26–27 October 2017, and on 8–9 October 2018 at Eötvös Loránd University, Budapest. The topics of both Seminars are relevant not only for the Early Modern philosophy but also for our contemporary philosophical attempts to find access to present reality: constructions of personal identity, and the multifarious relationship between theories and practices of natural right and the claims to live up to our natural emotions. When composing this volume, our aim was not to present a systematic survey of any of these areas of topics in Early Modern philosophy. Rather, our modest goal was to foster collaboration among researchers working in different countries and 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 both in early modern scholarship and in several related areas and disciplines.



ISBN 978-963-489-228-1



9 789634 892281