Sara Heinämaa
Martina Reuter

Editors

Psychology and Philosophy

*Inquiries into the Soul from
Late Scholasticism to
Contemporary Thought*

Springer

PSYCHOLOGY AND PHILOSOPHY

STUDIES IN THE HISTORY OF PHILOSOPHY OF MIND

Volume 8

Aims and Scope

The aim of the series is to foster historical research into the nature of thinking and the workings of the mind. The volumes address topics of intellectual history that would nowadays fall into different disciplines like philosophy of mind, philosophical psychology, artificial intelligence, cognitive science, etc. The monographs and collections of articles in the series are historically reliable as well as congenial to the contemporary reader. They provide original insights into central contemporary problems by looking at them in historical contexts, addressing issues like consciousness, representation and intentionality, mind and body, the self and the emotions. In this way, the books open up new perspectives for research on these topics.

PSYCHOLOGY AND PHILOSOPHY

INQUIRIES INTO THE SOUL FROM LATE SCHOLASTICISM TO CONTEMPORARY THOUGHT

Edited by

SARA HEINÄMAA

Academy of Finland and University of Helsinki, Finland

and

MARTINA REUTER

Academy of Finland and University of Helsinki, Finland

Editors
Dr. Sara Heinämaa
Helsinki Collegium for Advanced Studies
FI-00014 University of Helsinki
Finland
heinamaa@mappi.helsinki.fi

Dr. Martina Reuter
Department of Systematic Theology
FI-00014 University of Helsinki
Finland
martina.reuter@helsinki.fi

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Acknowledgements

Contents

Contributors

Tuomo Aho
Department of Philosophy, University of Helsinki

Pascal Engel
Department of Philosophy, University of Geneva

Edoardo Fugali
Department of Philosophy, University of Trieste

Prof. Gary Hatfield
Department of Philosophy, University of Pennsylvania

Sara Heinämaa
Academy of Finland, Helsinki Collegium for Advanced Studies,
University of Helsinki

Riku Juti
Department of Philosophy, University of Helsinki

Pekka Kärkkäinen
Department of Systematic Theology,
University of Helsinki

Henrik Lagerlund
Department of Philosophy, Talbot College, The University of Western Ontario

Sami Pihlström
Department of Social Sciences and Philosophy, University of Jyväskylä

Thomas Pink
Department of Philosophy, King's College London

Martina Reuter
Academy of Finland, Department of Systematic Theology, University of Helsinki

Camilla Serck-Hanssen
Department of Philosophy, University of Oslo

Frederick Stoutland
St. Olaf College (Minnesota) and University of Uppsala

Theo Verbeek
Department of History of Philosophy, Utrecht University

Mikko Yrjönsuuri
Department of Social Sciences and Philosophy, University of Jyväskylä

Dan Zahavi
Center for Subjectivity Research, Faculty of Humanities, University of Copenhagen

Introduction

Sara Heinämaa and Martina Reuter

Aristotle's *On the Soul* (*De anima*) is often presented as the first embryonic form of modern psychology. We are taught and trained to think that the Aristotelian concept of the soul covers the basic functions that today form the core areas of psychological research: perception, emotion, memory and intellect. On the other hand, some contemporary commentators are ready to argue that although Aristotle was interested in the directedness and reflexivity of perception and thought, his discourse on the soul does not conceptualize or systematize these structures as is done in modern philosophy of mind.[1] Thus, it may seem that Aristotle's *On the Soul* belongs more intimately to the history of psychology than to the history of the philosophy of mind.

It is less well known, and seldom emphasized, that Aristotle's *On the Soul* was used as a textbook in European universities well into the seventeenth century. For almost two thousand years Aristotelian psychology was considered an essential part of the curriculum of natural philosophy. Moreover, psychology, in parallel with physics and mathematics, was often studied as an example of the Aristotelian science of nature. So when Descartes' presented his account of scientific knowledge founded on the *ego cogitatio*, he addressed an audience well trained in classical Aristotelian psychology and its understanding of the principles and faculties of souls.

The idea that our knowledge of the soul is highly accurate and important was already present in Aristotle's *On the Soul*. The work starts with the following statement:

> We regard all knowledge as beautiful and valuable, but one kind more so than another, either in virtue of its accuracy, or because it relates to higher and more wonderful things. On both these counts it is reasonable to regard the inquiry concerning the soul as of the first importance.[2]

---

[1] For such arguments and alternative viewpoints, see Sara Heinämaa, Vili Lähteenmäki and Pauliina Remes, "Introduction," in Sara Heinämaa, Vili Lähteenmäki and Pauliina Remes (eds.), *Consciousness: From Perception to Reflection* (Dordrecht: Springer), 1–26.

[2] Aristoteles, *On the Soul*, 402a1–5, trans. W.S. Hett, Loeb Classical Library (Cambridge, Mass.: Harvard University Press, 1957). A more literal translation of Aristotle's text is provided by J.A. Smith: "Holding as we do that, while knowledge of any kind is a thing to be honoured and prized, one kind of it may, either by reason of its greater exactness or of a higher dignity and greater wonderfulness in its objects, be more honourable and precious than another, on both accounts we

As is well known, Descartes' meditations gave the soul, or mind, a new foundational role, both methodologically and metaphysically. Descartes grounded the whole system of the sciences on the mind's capacity to know itself as a thinking thing (*res cogitans*). The British empiricists problematized Descartes' deductive approach and argued that all knowledge, also all knowledge about the mind and its capacities, stems from experience and proceeds by association and abstraction. These thinkers developed subtle distinctions between different kinds of mental contents in order to defend this view, such as the distinctions between inner and outer sense, complex and simple ideas, ideas and impressions, and ideas of primary and secondary qualities. In 1780s Kant's transcendental turn radically changed the setting. Whereas Descartes still considered the mind to be a real thing, a substance, and thus knowable on the same metaphysical grounds as material substance, Kant strictly separated transcendental knowledge, including knowledge of cognitive capacities, from empirical knowledge of the world. This deeply reformed the epistemological question about our ability to know our own minds and proposed a whole new set of methodological and ontological problems.

Today, empirical psychology constitutes its own autonomous discipline; and conceptual studies of the nature of the mind, or consciousness, belong to subfields of the philosophical discipline: epistemology, philosophy of mind, history of philosophy and the theoretical parts of cognitive science. Psychology and philosophy are conceived and practiced as two distinct forms of knowledge, one empirical and one conceptual or transcendental.

The present volume investigates the Western tradition of philosophical psychology and its relation to scientific psychology from the beginning of the sixteenth century to the present. The aim is to question some deep-seated convictions about the nature of these two sciences and their objects of study. Our interest in the history of thought is twofold: We turn back in order to learn from our ancestors and in order to question our own habitual notions about the mind, the soul and the psyche, and their mutual relations. Instead of simply buttressing the received view of one unified line of development – having its starting point in Aristotle and two major turning points in Descartes and Kant – the articles of the book present several parallel developments and a series of incremental changes. This does not mean that the profound impact of Descartes and Kant is overlooked, quite the contrary. The aim is to deepen our understanding of their influence by unveiling less well known lines of development, such as the post-Kantian Aristotelianism of the nineteenth century. We find in these developments new inventive accounts of the unity of the soul as well as the unificatory and differentiating aspects of perception.

In the first chapter, Gary Hatfield studies the origin of contemporary psychology and philosophy of mind from the beginning of the seventeenth century to the twentieth century. He examines the early modern conception of "the place of the

---

should naturally be led to place in the front rank the study of the soul," trans. J.A. Smith, in *The Complete Works of Aristotle*, vol. 1, ed. Jonathan Barnes (Princeton: Princeton University Press, 1984).

mind in nature," challenging two recurring notions about Descartes' understanding of the mind: the notion that Descartes's dualism located mind outside nature and the assumption that his mind–body dualism equates a division between the psychological and the non-psychological. Hatfield argues that Descartes's attempt to explain sensory phenomena required reference both to mind and to matter. Moreover, he uses his close reading of Descartes as a basis for the rehabilitation of the two most problematic aspects of modern psychology: introspection and behaviorist modes of explanation. Hatfield distinguishes between the different tasks attributed to introspection and argues that introspection can serve as a valuable tool for psychological knowledge if its function is restricted to the reporting of mental occurrences, without any commitment to the transparency of the mind to itself. In an analogous way, Hatfield shows that behaviorism is problematic as a general methodology but valuable as a restricted strategy for the explanation of specific cases of psychological phenomena.

Chapters 2 and 3 turn back to investigate the Aristotelian heritage. They show how Aristoteles' works were received, interpreted and developed in the Renaissance and early modern thinking. In Chapter 2, Pekka Kärkkäinen and Henrik Lagerlund examine the contents and methods of philosophical psychology as it was taught in the sixteenth century at the universities of Erfurt, Padua and Bologna. Contrary to the received notion of the Renaissance, Kärkkäinen and Lagerlund argue that philosophical psychology of the sixteenth century was tightly bound to the Aristotelian tradition. The chapter also discloses historical differences between the Erfurt school and the North-Italian centres of learning. Philosophical psychology in Erfurt had a strong connection to the Buridanian tradition, which coloured discussions of the metaphysical nature of the soul as well as disputes about universal realism versus nominalism. In Padua and Bologna, on the other hand, metaphysical debates on the nature of the soul culminated in a dispute about the limits of natural reason. Pietro Pomponazzi argued that the immortality of the human soul cannot be proven by natural reason. He defended the position that the question must be left open: the soul can neither be proved mortal nor immortal.

In Chapter 3, Tuomo Aho studies the status of psychological knowledge as it was understood by the leading scholastic philosophers of the sixteenth century. Aho focuses his discussion on the very beginning of Aristotle's *On the Soul*, the first sections of the first chapter, which consists of some important methodological remarks. He examines a number of scholastic Aristotle-commentaries, in particular two great Jesuit works, the so-called Coimbra-commentary and Francisco Suárez's more original *De anima*. Aho's inquiry demonstrates that most sixteenth-century commentators had a clear position on the question concerning the status of psychology as a science. The chapter also suggests that in the context of this "second scholasticism," the Aristotelian theory on the soul was studied and used as an exemplary science. In these discussions, we can already find insights that later became systematized in Descartes' discourse on the mind.

Together these two chapters give us a new perspective on sixteenth-century centres of learning by illuminating the richness and versatility of Aristotelian psychology. What we find is not one unitary theory but a flexible approach which

included several disparate conceptions of the soul, its nature and functions, and its relations to other natural and supra-natural phenomena. Thus, we come to see that Aristotelian psychology is not in any simple oppositional relation to the modern disciplines of psychology, epistemology, and philosophy of mind. With regard to the science of the soul, we cannot conclude, as is often done, that modern science was born out of the breakdown of Aristotelianism. On the contrary, it originated and developed in an Aristotelian environment.

In Chapter 4, Mikko Yrjönsuuri continues the investigation of Cartesianism by asking what it would mean to talk about psychology in Descartes' terms. Yrjönsuuri argues that the possibility of Cartesian psychology depends as much on Descartes' concepts of science and human knowledge as on his concepts of the soul and the thinking self. In his interpretation, Descartes' epistemological framework does not allow to formulate psychological questions in the Aristotelian way, because the Cartesian distinction between the psychic and the physical differs from the Aristotelian one. Neither does Cartesianism include questions of the type that we encounter in contemporary psychology or philosophy of mind. Yrjönsuuri agrees with Hatfield in arguing that Descartes formulates psychological problems both in terms of thinking (mind) and in terms of extension (matter), but he argues that Descartes' concept of the intellectual soul cannot be naturalized in the same way as his concepts of the sensing soul and the perceiving soul.

Chapter 5 studies Descartes' rationalist successor and critic, Benedict Spinoza. Theo Verbeek clarifies the relation between three modes of thought that are central to Spinoza's account of our knowledge and understanding of the human mind: the concepts of intellect, imagination and reason. For Spinoza, the highest form of cognition is the intuitive understanding achieved by the intellect. Reason-based knowledge remains secondary in relation to intuitive knowledge and provides general and abstract ideas, which do not denote any real existing entities but merely have an instrumental role in the explanation of the relations between such entities. Verbeek shows that in his final analysis, Spinoza conceives these general ideas as *entia imaginationis*, products of imagination, and he concludes that Spinoza's notion of reason is closer to the notion of imagination than to that of the intellect.

Early modern reconceptions of the soul also had far-reaching consequences in the areas of moral and political philosophy. Resting on the philosophy of Thomas Hobbes, the British empiricists elaborated and developed the idea that human action is grounded on beliefs and motivated by desires. In Chapter 6, Thomas Pink shows how this new focus on motives and desires related to traditional scholastic explanations, which were informed by the Aristotelian conception of practical reason. Pink illuminates the emergence of the so-called belief-desire model of the soul and shows how it affected conceptions of moral responsibility. He contrasts the theories of Samuel Pufendorf and John Locke, claiming that the former remains close to the scholastic theory of moral obligation adopted and developed by Suárez, whereas the latter breaks with this tradition and introduces a new conception of agency based on beliefs and desires.

Enlightenment philosophy witnessed the first wave of a strictly empirical philosophy of mind with inductive methods. Riku Juti examines the strengths and weak-

nesses of inductivism in Thomas Reid's attempt to develop a scientific philosophy of the mind. He argues that Reid's concept of induction is Newtonian and should be understood as a means of discovery rather than as a means of justification. Moreover, Juti suggests that Reid's distinction between things "in the mind" and things "external to the mind" is not ontological but relational or intentional. The chapter ends with an analysis of Reid's famous principles of common sense. Juti argues that Reid's principles come close to the general propositions postulated by medieval philosophers and that Reid can be correctly understood only if one remembers that he was a pre-Kantian philosopher.

The eighth chapter discusses the Kantian turn in the philosophy of consciousness and experience. Camilla Serck-Hanssen explicates Kant's concepts of apperception and inner sense, and studies them in relation to his sceptical attitude towards the idea that psychology could be regarded as an empirical science. In Kant's terminology, inner sense is the consciousness of the events that take place within the mind whereas apperception is the reflexive consciousness of one's own spontaneous acts. These two concepts generate two different questions concerning the relation between consciousness and nature. On the one hand, there is the question of how inner nature or mental nature is related to physical nature; on the other hand, there is the question of how spontaneity is related to the whole of nature – inner nature as well as outer nature. Serck-Hanssen argues that Kant's answer to the first question is closely related to his pessimism regarding the status of psychology as an empirical science. The latter question is more far-reaching and it also has significance for Kant's moral philosophy and aesthetics.

Kantian philosophy introduced new criteria for scientific knowledge and philosophical thinking. In Chapter 9, Martina Reuter studies how the physiognomic theory of Johann Caspar Lavater exemplifies the modern division between science and pseudo-science. Lavater's aim was to identify human character traits and dispositions by making observations on a person's physical appearance. Reuter argues that Kant's distinction between science and art helps to understand and explicate the pseudo-scientific nature of Lavater's theory. The chapter concludes by showing that Mary Wollstonecraft adapted some of Lavater's ideas. Wollstonecraft was an early feminist interested in the cultivation of human character. She shared Lavater's assumption that there is a correspondence between bodily constitution and character, but she remained true to the Enlightenment spirit by emphasizing the role of education and by questioning Lavater's preference for natural necessity over human freedom.

Nineteenth-century romanticism and idealism witnessed a revival of the Aristotelian tradition. The Aristotelianism of the nineteenth-century Aristotelians, such as Friedrich Adolf Trendelenburg and Franz Brentano, was profoundly informed by Kant's conception of knowledge. Since the beginning of the century, Kant's distinction between critical philosophy and positive science had guided inquiries into mind and consciousness. In Chapter 10, Eduardo Fugali argues that Trendelenburg and Brentano returned to Aristotle's concepts in order to develop a philosophy of mind which would not depend on empirical psychology but would secure its legitimacy by

philosophical insight into the essence of the psyche. The importance of Aristotelian concepts for these thinkers is manifest both in the topics that they discussed and the methods that they used. Fugali shows how Aristotelian and Kantian influences affected Trendelenburg's and Brentano's views of the epistemological status of psychology and its demarcation from and dependence on philosophy.

The unclear nature of introspection as a method of inquiring into the psyche or the mind burdened all philosophical movements during the nineteenth and early twentieth century. Sami Pihlström studies how William James's pragmatic metaphysics of the mind tried to overcome the idea of a self-enclosed mind and account for our relations to other minds. He emphasizes James's view of the mind as an active and purpose-oriented organizing principle which structures our lifeworld. James's pragmatism is usually classified among empiricist and associationist philosophies of mind, but Pihlström illuminates its Kantian features and compares it to twentieth-century phenomenology. James's philosophy is also often characterized as individualistic and even accused of solipsism. Pihlström argues that James's position can be properly understood only if we realize that for James the problem of other minds was not merely metaphysical or epistemological, but also, perhaps even primarily, ethical.

In the twelfth chapter, Pascal Engel discusses the spiritualist tradition that developed in the eighteenth and nineteenth centuries in France. Engel examines the spiritualist movement and its different variations, from Maine de Biran to Henri Bergson. He shows that the movement involved several different types of inquiry: psychological and mainly introspective analyses as well as metaphysical investigations of selfhood, being and time. Engel takes a critical stance and argues that the speculative method employed by the spiritualists distanced them from experimental psychology and resulted in an unhappy attempt to derive metaphysics from introspection. Engel compares the spiritualist tradition with Brentano's descriptive psychology and argues that Brentano's approach is more promising, because his psychology was embedded in a metaphysics which was independent of introspective investigations. Brentano was liberated from introspectionism by his analytical approach and his theory of parts and wholes.

During the first half of the twentieth century, phenomenologists developed an influential reformulation of the Kantian distinction between critical-transcendental philosophy and naturalistic philosophy. This development was also influenced by Brentano's revival of Aristotelianism and by the epistemological radicalism of Descartes and Hume. In Chapter 13, Dan Zahavi clarifies Edmund Husserl's idea of a transcendental science of consciousness and its operations, and shows how Husserl's transcendentalism differs from Kant's. Husserl distinguished between three different approaches to mental operations: empirical psychology, eidetic psychology, and transcendental phenomenology. According to Husserl, the first two deal with mental processes as real occurrences: empirical psychology studies psychic processes as part of the causal network of nature, and eidetic psychology clarifies the psyche as an autonomous system, in abstraction from the causal connections that are determining within the naturalistic stance. Transcendental phenomenology

has the fundamental epistemological and methodological task of providing the on-tological ground for both these sciences by explicating the sense of the psyche.

In Chapter 14, Sara Heinämaa studies the development of Gestalt psychology and its relation to phenomenological philosophy. Both were influenced by Brentano's theory of parts and wholes, but whereas the Gestalt-theorists focused their inquiries primarily on the structures of sense-perception and sensation, the phenomenol-ogists used Brentano's distinctions to establish a comprehensive philosophy of consciousness, which would cover all types of intentional acts and objects, from sense-perceptions to memories, phantasies, cognitions, emotions and volitions. Heinämaa points out that both Gestalt-theorists and phenomenologists struggled to liberate their accounts of perception from atomistic and dualistic prejudices. Despite this shared interest, the two approaches remain far apart in their principles, because the former studies the mind as a worldly entity or process whereas the latter aims at establishing a transcendental foundation for all mundane knowledge, psychology as well as physics, through the methods of reduction. Heinämaa shows that the transcendental-phenomenological reduction is central, not just to Husserl's classical phenomenology, but equally to Maurice Merleau-Ponty's phenomenology of em-bodiment. Thus, Merleau-Ponty's engagement with Gestalt-psychology is not just constructive but also critical.

The last, fifteenth chapter demonstrates how Wittgenstein's inquiries into lan-guage influenced twentieth-century philosophy of mind. Frederick Stoutland argues that Wittgenstein's most significant contributions were his rejection of the Cartesian distinction between the mental and the physical and his emphasis on the primacy of action over intellection. Stoutland attacks the different versions of contem-porary naturalism, from Jerry Fodor's causalist theory of mind to Paul Church-land's eliminativism. He argues that both views assume a distinction between the physical and the mental; both identify physical nature with a physicalist system; and both understand the mental as a realm of inner, non-public, entities. Stout-land presents Donald Davidson's anomalism as a viable alternative to Fodor's and Churchland's naturalism. He emphasizes that Davidson attacked physicalism as well as the mental–physical distinction which underlies it. For Stoutland, Davidson is a neo-Wittgensteinian thinker and the most promising developer of Wittgenstein's anti-Cartesianism.

Chapter 1
Psychology in Philosophy: Historical Perspectives

Gary Hatfield

Abstract The chapter examines some common assumptions regarding the shape of the history of theories of mind. It questions the conception that the Scientific Revolution resulted in placing the mind "outside of nature." During the seventeenth century, the followers of Descartes routinely placed study of the mind, or, at least, mind–body interaction, within "physics" considered as a science of nature in general (and so including physics in the narrow sense, biology, and psychology). By the end of the eighteenth century, many authors treated the psychological and the physical as distinct areas of investigation (with points of contact, to be sure), both amenable to natural scientific study. Another assumption is that, prior to behaviorism, the notion of "the psychological" was equivalent to "the mental." And yet, during the seventeenth century psychological phenomena were assigned to both sides of the mind–body divide: the Cartesians and their allies offered (nonmentalistic) mechanistic explanations for the sensory and behavioral capacities that human beings share with animals, while also holding that some aspects of mental life (including consciousness, free will, conceptual language, and general intelligence) require an immaterial mind. Finally, from the late eighteenth century onward, some have maintained that if mental states are reached only by introspection then they cannot be studied objectively. The chapter argues that the standard version of this objection misunderstands the nature of introspection, and it mounts a reply founded on Wundt's methodological discussions and making reference to the perceptual investigations of the Gestalt psychologists.

Keywords Mind–body relation · mind and nature · introspection · Descartes · Wundt

The history of theories of mind has most often been studied by examining the theories and positions of individual thinkers. Scholars (especially) of early modern philosophy have examined recurrent themes, such as the mind–body problem and the perception of primary and secondary qualities, with special attention to Descartes or

G. Hatfield
Department of Philosophy, University of Pennsylvania

Locke. Recent studies have expanded the range of topics to include the physiological basis of Descartes' psychological theories, the passions and emotions in a range of authors, and the scholastic background to early modern theories of mind.[1] These studies belong to a resurgence of work in the history of early modern philosophy more generally, fueled by a methodology of historical contextualization that gives proper attention to the relations between philosophical disciplines such as metaphysics and natural philosophy and other intellectual pursuits in the early modern period, such as the mathematical and empirical sciences, or theology. Historians of philosophy have given less attention to history of mind and psychology in the nineteenth and early twentieth centuries, though that is changing.

Comparatively little systematic attention has been given to the history of mind beyond the treatment of specific problems and themes in individual authors (or pairs of authors). That is, few scholars have devoted critical attention to what I call the "shape" of the history of theories of mind: to the narrative and explanatory structures by which scholars link author to author, period to period, and topic to topic. More generally, historians of philosophy and psychology have not attended to the various disciplinary loci of psychology and theories of mind in the modern period. It is only now becoming widely appreciated that, in the orthodox Aristotelian scheme, the topics of *De anima* (or *Peri psychēs*) fell under the rubric of physics, understood as the science of nature in general (and so including subject matters that we would now describe as physical, biological, and psychological). It is even less well known that Descartes and his followers placed mind–body interaction within physics, which they continued to treat as the science of nature in general, and that some Cartesians, and many eighteenth-century thinkers, regarded the immaterial mind to be a natural entity, a thing in nature.

Plenty of historical and philosophical work will be needed in order to understand past theories of mind in their historical contexts and to forge well-conceived narrative frameworks. Concurrently, we should hope that this historical work, by shedding light on the origin and development of present-day positions and questions, will have significant philosophical payoffs for present-day work in philosophy of mind. It is in fact inevitable that contemporary theory is reciprocally tied to historical work, so that each conditions and can illuminate the other. Our current understanding of the "philosophical problems" of mind typically embodies a tacit conception of how those problems arose. These tacit conceptions amount to unexamined historical theses about the origin and structure of the contemporary problems, which means that new historical understanding can yield new conceptions of the structure of contemporary problems. At the same time, our contemporary understanding of these problems conditions the thoughts of historians in the very doing of history. Thus, historians may well imbibe undetected historical theses while being trained in or engaging contemporary work in the philosophy of mind, through historical

---

[1] Two recent collections, ranging from antiquity to modernity: Tim Crane and Sarah Patterson (eds.), *History of the Mind–Body Problem* (London: Routledge, 2000), and John P. Wright and Paul Potter (eds.), *Psyche and Soma: Physicians and Metaphysicians on the Mind–Body Problem from Antiquity to Enlightenment* (Oxford: Clarendon Press, 2000).

theses that are implicit in the accepted problem space. These theses extend not only to "received opinions" about individual authors but also to the shape of the history of theories of mind more generally. Such influence is inevitable, since historians of philosophy are typically trained as philosophers before the time when they become reflective about methodology in the history of philosophy.

In this chapter I will examine two historical assumptions about the shape of the history of mind. The assumptions pertain to the "place of the mind in nature" in early modern philosophy, and to the introspective method as employed in late nineteenth- and early twentieth-century philosophy and psychology. I combine these apparently disparate topics because I believe that each reveals the lingering effects of attitudes toward mind and psychology that became widespread when mentalistic psychology separated from philosophy in the early twentieth century, and when, in subsequent decades, behaviorism spread from psychology to philosophy and became entrenched (roughly, 1930–1970).[2] The relations and effects of these two historical assumptions are further specified in Section 1.1.

The legacy of behaviorism deserves special attention. Behaviorist attitudes, by engendering a conception of the defects and defeat of introspective methods in psychology, have been able to influence current philosophical attitudes and to color the portrayals of past philosophical positions. Of course, all or most philosophers who have been influenced in these ways would now repudiate behaviorism as a philosophical or methodological position, and many would consider themselves to be initiators or heirs of the "cognitive revolution". Nonetheless, even after behaviorism's demise as an avowed position, it retains an unnoticed influence. More generally, the positive legacy of behaviorism has not been adequately appraised. Indeed, one of the historical episodes analyzed herein provides evidence that, when viewed from a fresh standpoint, behaviorist concepts can provide useful analytical tools for the history of theories of mind, including the relations between the mental and the psychological.

The historical interpretations I offer regarding the place of mind in nature, the relations between the mental and the psychological, and the alleged demise of introspection are put forward in a revisionist spirit. The aim of the revision is to take a fresh look at the sources. The need for it stems from the way that early twentieth-century attitudes have conditioned much historical work.

1.1 Hindrances to Histories of Theories of Mind

During the middle part of the twentieth century, "mind" was a dirty word in Anglophone philosophy. Two historical forces combined to make this so. First, in the latter nineteenth and early twentieth centuries, the thesis became widely accepted

---

[2] I do not mean to suggest that behaviorism arose in psychology without any prior philosophical influences. On the origin and development of behaviorism, see Gary Hatfield, "Behaviorism and Psychology," in Thomas Baldwin (ed.), *Cambridge History of Philosophy, 1870–1945* (Cambridge: Cambridge University Press, 2003), 640–48, and the literature cited therein.

that psychological description of mental processes, and any purportedly "scientific" psychology of mind, is irrelevant for epistemology and for philosophy more generally. This attitude became embodied in the notion of a "psychologistic fallacy" with respect to epistemology and logic. According to this "fallacy," it is a category error to think that an account of psychological processes could be relevant to an assessment of the logical adequacy or epistemic warrant of thoughts. Acceptance of the legitimacy of this outlook came to condition the history of early modern theories of mind. For example, it was assumed that reference to cognitive faculties or to a theory of mind in evaluating the possibility and limits of human knowledge were instances of this fallacy. Descartes' talk of innate ideas was assimilated to a simple thesis asserting psychological innateness, and was dismissed as epistemically irrelevant, since, in truth, innateness of itself bestows no epistemic warrant.[3]

The second factor arose through the behaviorism of John Broadus Watson, which was vigorously antimental, proposing to do without any mentalistic notions whatsoever. Watson's behaviorism caught the attention of philosophers as diverse as Bertrand Russell, Rudolf Carnap, and W.V.O. Quine – though Russell, unlike the others, didn't accept it wholeheartedly.[4] The behaviorist credo that, to be scientific, psychology must reject not only "mind" as a substantial entity but also mentalistic concepts and descriptive vocabulary, became widespread. Ultimately, psychology was to be reducible to physics and chemistry. Many philosophers, including Quine and his followers, concluded that mentalistic talk was empty.[5]

These two trends were associated with an historical thesis about the place of the mind in nature, and with a methodological thesis about the relation of mental phenomena to natural scientific investigation. The historical thesis was that, as the mechanistic view of nature came to dominate the seventeenth century, it relegated mind to a place "outside" nature. Nature was equated with the system of

---

[3] On the historiographical influence of the psychologistic fallacy, see Gary Hatfield "The Workings of the Intellect: Mind and Psychology," in Patricia Easton (ed.), *Logic and the Workings of the Mind: The Logic of Ideas and Faculty Psychology in Early Modern Philosophy*, North American Kant Society Publications vol. 5 (Atascadero: Ridgeview Publishing Co., 1997), 21–45, and "Epistemology and Science in the Image of Modern Philosophy: Rorty on Descartes and Locke," in Juliet Floyd and Sanford Shieh (eds.), *Future Pasts: Reflections on the History and Nature of Analytic Philosophy* (New York: Oxford University Press, 2001), 393–413. On the history and sociology of psychologism, see Martin Kusch, *Psychologism: A Case Study in the Sociology of Philosophical Knowledge* (London: Routledge, 1995).

[4] Bertrand Russell, *Analysis of Mind* (London: Allen & Unwin, 1921); Rudolf Carnap, "Psychologie in physikalischer Sprache," *Erkenntnis* 3 (1932), 107–42, trans. G. Schick, "Psychology in Physical Language," in A. J. Ayer (ed.), *Logical Positivism* (New York: Free Press, 1959), 165–97; W.V.O. Quine, *The Roots of Reference* (LaSalle, Ill.: Open Court, 1973). Some versions of behaviorism, including E.C. Tolman's, did not reject all mentalistic concepts; see Hatfield, "Behaviorism and Psychology".

[5] P.S. Churchland, *Neurophilosophy: Toward a Unified Science of the Mind-Brain* (Cambridge: The MIT Press, 1986), is an example. Daniel Dennett, *The Intentional Stance* (Cambridge: The MIT Press, 1987), while not holding that mentalistic notions are completely empty, treats them as useful instrumental fictions, ultimately to be cashed out by explanations given from the "physical stance".

material bodies or their constituents, or perhaps with the mathematically describable properties of such bodies or their constituents. Barring the reductivist conclusion – which few early moderns were willing to accept – that matter can think (and hence that thought is material and so "within" nature), thought and the mental were thereby excluded from nature. The conclusion that mind lacks a place in nature was reinforced during behaviorist times, when references to mind might be treated not merely as scientifically intractable but also as equivalent to a belief in ghosts.[6]

The methodological thesis is that mental phenomena are resistant to the techniques of natural science: unlike material entities and processes, minds and/or mental phenomena cannot be studied and measured objectively. Kant raised this objection in his own way, by denying that psychology could be a mathematical science (according to his peculiar definition of "mathematical science").[7] Watson and other behaviorists made this methodological thesis a mainstay in their arguments against introspective psychology. In a limited form, the objection is repeated in much recent philosophy of mind, which disparages the notion of introspection and rejects the possibility of introspective access to peculiarly mental phenomena.

This methodological thesis reinforces, and is reinforced by, the historical thesis. One might be led to assume that the mind was excluded from nature *because* it was not susceptible to the mechanizing and mathematizing tendencies of the mechanical philosophy. One might further assume, and many have assumed, that behaviorism actually revealed that introspective methods in psychology are completely bankrupt. Hence, in the name of rigor, the new cognitivist approaches in psychology and philosophy of mind would have to exclude phenomenality or introspectable phenomenal content.

All the same, these new cognitivist approaches have inspired some revised attitudes toward the relation between epistemology and psychology (as in reliabilist epistemology),[8] and they have sought to rehabilitate mentalistic concepts such as belief, desire, and intentionality. But the rehabilitation has been selective and conditional. The most widely accepted view about mentalistic concepts is that, to retain legitimacy, they must reduce to allegedly physicalistic notions such as information, or perhaps to basic physics itself. To paraphrase a comment from Jerry Fodor a few years ago: "If intentionality and phenomenality are real, they must really be something else".[9] That is, the "natural" (and hence the explanatorily legitimate) is

---

[6] J.B. Watson, "Psychology as the Behaviorist Views It," *Psychological Review* 20 (1913), 158–77; B.F. Skinner, "The Operational Analysis of Behavioral Terms," *Psychological Review* 52 (1945), 270–77, and "Behaviorism at Fifty," *Science* 140 (1963), 951–58.

[7] On Kant's relation to psychology, his criticism of introspection, and his peculiar criteria for applying the label "natural science," see Hatfield, "Empirical, Rational, and Transcendental Psychology: Psychology as Science and as Philosophy," in Paul Guyer (ed.), *Cambridge Companion to Kant* (Cambridge: Cambridge University Press, 1992), 200–27.

[8] Alvin Goldman, *Epistemology and Cognition* (Cambridge: Harvard University Press, 1986).

[9] Fodor actually said: "If aboutness is real, it must be really something else" (*Psychosemantics: The Problem of Meaning in the Philosophy of Mind*, Cambridge: The MIT Press, 1987, 97). He was suggesting that realism about intentionality ought to be paired with reduction (to physical states). In chapter 4, he expressed skepticism about teleological approaches to naturalizing content.

co-extensive with the "physical" (as described using the terms of physical science), or perhaps with physics plus evolutionary biology (as many theorists claim, though not Fodor, who doubts that evolution can help here).[10] Although some voices of protest are heard,[11] the dominant outlook remains Quinean: mentalistic terms may be used for convenience, but when push comes to shove, they must be conceptually reducible to physics and biology. From this perspective, phenomenality is difficult to assimilate into the natural order, unless it proves reducible. This outlook has engendered its own historical assumptions about the objectives for using introspection in studying sense perception.

1.2 Descartes' Naturalism About the Mental

Descartes supposedly was a leader in excluding mind from nature through his mind–body dualism. He is commonly understood to have divided the states and processes of human beings between those that depend on the mind alone, which are therefore mental (and outside nature), and those that belong to the body alone, which are therefore mechanical, nonmental, and hence nonpsychological. In this scheme, he had to acknowledge some anomalous "left over" mind–body interactions or relations that he was unable to account for metaphysically or to treat coherently in his natural philosophy.[12]

I would like to challenge both the thesis that Descartes and his followers placed the mind outside nature (in this section) and the thesis that his mind–body dualism constituted a division between the psychological and the nonpsychological (in Section 1.3.). In their discussions of natural philosophy and the principles of the natural world, Descartes and some of his followers tacitly or explicitly placed the mind, or at least mind–body interaction, within the domain of nature. Cartesians generally treated mind–body interaction within physics or natural philosophy, and in doing so they implied or asserted that the interaction occurs according to laws or rules that might be empirically investigated. Further, those Cartesians who explicitly placed the immaterial mind under the rubric of "nature" treated it as an extant force or agency whose processes and effects were subject to investigation within natural philosophy. They did not intend to imply any sort of materialistic reduction, nor to

---

[10] The best developed attempts to use biofunctional notions to naturalize intentionality stem from Fred Dretske, *Explaining Behavior: Reasons in a World of Causes* (Cambridge: The MIT Press, 1988) and *Naturalizing the Mind* (Cambridge: The MIT Press, 1995). It should be noted that Dretske begins from a notion of "information" as a mind-independent physical state that naturally possesses low-grade intentionality, from which cognitive-grade intentionality evolves; for Dretske's concept of information, see his *Knowledge and the Flow of Information* (Cambridge: The MIT Press, 1981).

[11] Gary Hatfield, *The Natural and the Normative: Theories of Spatial Perception from Kant to Helmholtz* (Cambridge: The MIT Press, 1990), Ch. 7; Galen Strawson, *Mental Reality* (Cambridge: The MIT Press, 1994).

[12] On Descartes' mind–body interaction as anomalous, see Lilli Alanen, *Descartes's Concept of Mind* (Cambridge: Harvard University Press, 2003).

require that the mind's own operations (as opposed to the mental and bodily effects of mind–body interaction) occur deterministically. If there were such a thing as spontaneous free will, that, too, might be a property of a natural being, a property of an immaterial substance existing in nature.

Having elsewhere presented these claims in some detail,[13] I will here examine and interpret some quotations from Descartes that seem to challenge this thesis, but which, in larger context, support it. I will then discuss the works of two Cartesians, before characterizing some eighteenth-century developments.

Some of Descartes' own words seem to restrict the physical and the natural to the material. Most noteworthy is a passage from the *Principles of Philosophy* (Part II) in which Descartes says, "I do not accept or desire in Physics any other principles than in Geometry or abstract Mathematics; because all the phenomena of nature are explained thereby" (Pr II.64).[14] Previously, he had asserted that the nature of body consists in extension (Pr II.4), which he now affirms to be the object of geometry. He further claims that, from the "divisions, shapes, and movements" of material substance, "all Natural Phenomena" can be explained, and he then concludes that "no other principles of Physics should be accepted, or even desired" (Pr II.64). This statement seems clear enough, but a reader of Descartes must be wary, for he sometimes sets out a categorical conclusion only to adjust or modify it a page or two further on.

Indeed, in the very next article, Descartes says that, while the truth of these principles about material things should not be doubted (for the principles were discovered by the "light of reason"), it nonetheless remains to be seen "whether we are able to explain all the phenomena of nature by these principles alone" (Pr III.1). Throughout Parts III and IV, the *Principles* move along, explaining various cosmological, astronomical, planetary, geological, gravitational, thermal, tidal, mineralogical, pyrological, chemical, vitreous, and magnetic phenomena. Thus far, the principles of material things have been adequate, on the supposition that *motion* is a given (in Part II he described the causal basis of motion, which rests in God). But as he reaches the point where the two unfinished parts of the *Principles* would have begun (these parts were to have covered plants and animals and man), his outlook changes. He acknowledges that he has thus far "described this Earth, and indeed this whole visible world, as a machine, considering nothing in it except figures and motions" (Pr IV.188). Now he must explain other things that "our senses show us," including "colors, odors, sounds, and similar things". Without an explanation of these phenomena, Descartes concedes, he would "seem to have omitted the principal part of the explanation of natural things" (Pr IV.188).

---

[13] Gary Hatfield, "Descartes' Naturalism about the Mental," in Stephen Gaukroger, John Schuster, and John Sutton (eds.), *Descartes' Natural Philosophy* (London: Routledge, 2000), 630–58.

[14] Citations to the *Principles of Philosophy*, trans. V.R. Miller and R.P. Miller (Dordrecht: Reidel, 1983) are given in the text using the abbreviation "Pr" followed by part and articles numbers (or, for the front matter, lowercase roman numerals). Other references to Descartes' works are to the standard edition of Descartes' *Oeuvres*, ed. by Charles Adam and Paul Tannery, new edn. (Paris: Vrin/C.N.R.S., 1964–76), cited as "AT" followed by volume and page numbers (my translations).

In order to explain these phenomena, Descartes employs the concept of "sensations" and their causes. These explanations introduce into his world the "human soul" or mind, its seat in the brain, and the principles of mind–body interaction. The sensory qualities themselves, insofar as they reside in bodies, are nothing but arrangements of material particles that affect sensory media and hence the sense organs in various ways, thereby setting up motions in the nerves and brain. He goes on:

> [...] the movements which are thus excited in the brain by the nerves affect the soul or mind (which is closely joined to the brain), in diverse ways, according to the diversity of these movements. And these diverse states or thoughts of the mind, following immediately from these movements, are called the perceptions of the senses, or, as we commonly say, sensations. (Pr IV.189).

To explain sensory phenomena, both matter and mind are required. Hence, a physics of matter alone would be incomplete, because it would fail to explain these phenomena.

Descartes sets out various principles of mind–body interaction, including the principle that the differences among sensations (both external and "internal," to include passions and appetites) depend both on anatomically distinct nerves along with related brain states (some for hearing, some for vision, etc.), and on variations in the activity of any one of these nerves or brain states (Pr IV.189–98; see also AT 6:113–14, 130–41). Some nerves, if stimulated, give sensations of color, others of sound, others of odor, and so on (a principle known later as the Law of Specific Nerve Energies).[15] Within a given nerve group (supporting vision, or sound, or odor, etc.), the various ways in which the nerves are stimulated cause sensations in the mind that vary in intensity and quality (producing sensations of, say, red or blue, or high or low pitch, etc., with various intensities). If stimulated one way, these nerves cause one sensation, if another way, another sensation, all according to regular rules, or what his followers would call "laws". Descartes used empirical arguments to establish such regularities, seeking to correlate inferred brain states (inferred by working "in" from known physical stimuli) with experienced sensations (Pr IV.196–98; also, AT 6:131–33).

This apparent change of mind between Parts II and IV in fact reflected Descartes' considered view. In the Author's Letter to the French translation of the *Principles*, published in 1647, he included within physics the study of "the nature of man" (Pr xxv–xxvi), which involved mind, body, and their union (AT 3:508; 7:88). Moreover, in the *Passions of the Soul*, Descartes claimed to study the passions (which he defined as effects of the body on the mind) not as "an Orator, or even as a Moral Philosopher, but as a Physicist" (AT 11:326). In that work he was not merely studying the material causes of the passions, but he was examining the functions and causal implications of the passions themselves (which are intrinsically mental

---

[15] Hatfield, "Descartes' Naturalism about the Mental," 641–42.

states), all as a "physicist" or someone who studies the phenomena of nature.[16] Furthermore, several of Descartes' followers included study of the mind, at least in its relation to the body, within the domain of physics (as the science of nature). Pierre Sylvain Regis, *Système de philosophie* (1691), put the study of mind as an immaterial being in metaphysics only, but he placed in physics the mind–body union and its "laws," including those governing the production of sensations.[17] Antoine Le Grand, *Entire Body of Philosophy* (1694), placed the study of the soul as an immaterial being under "pneumatics" in metaphysics. He placed the "Science of Man" under "special physics" in the section on "living creatures". He divided the Science of Man into two parts: "Man Considered with Relation to his Body" included the bodily causes of sense, internal sensations, imagination, memory, and waking and sleep; "Man Considered in the other Part, The Mind" covered the distinctness of mind and body and their union, innate ideas, mental faculties, and the passions. For Le Grand, at least, mind itself fell under both metaphysics and physics.[18]

Descartes and his followers were not alone in placing the mind within nature, or in thinking that aspects of the mental, such as the laws of mind–body interaction, could be studied empirically. Throughout the eighteenth century, the mental and the psychological were assigned various disciplinary loci. Some placed the mind outside nature, in metaphysics (as did the Dutch Newtonian Gravesande). Others attempted to bring mental states and processes into the natural world, which, on their conception, was the world of matter (Diderot, d'Holbach, Priestley). Even if we construe these last efforts as attempts to "reduce the mental to the physical," these philosophers were *not* seeking to reduce the mental away. Still others, who accepted dualism, nonetheless held that the soul could be made the object of empirical study, and these were by far the majority in the first half of the century. A common position by the end of the eighteenth century was what C.C.E. Schmid called "empirical dualism".[19] An adherent of this position could remain agnostic on the ontology of mind and body while asserting that the mental and the physical are distinct domains of phenomena, each of which can be studied through observation and experiment, as can the relations between them. Those tacitly or explicitly adopting this attitude included Johann Gottlob Krüger, Guillaume-Lambert Godart, Charles Bonnet, David Hartley, and Schmid. Adam Ferguson placed the

---

[16] On Descartes' treatment of the passions as natural phenomena, see Daisie Radner, "The Function of the Passions," in B. Williston, and A. Gombay (eds.), *Passion and Virtue in Descartes* (New York: Humanity Books, 2003), 175–87, and Gary Hatfield, "The *Passions of the Soul* and Descartes' Machine Psychology," *Studies in History and Philosophy of Science* 38 (2007), 1–35.

[17] Pierre Sylvain Regis (1632–1707), *Système de philosophie: contenant la logique, metaphysique, physique & morale*, 7 vols. (Lyon: Anisson, Posuel & Rigaud, 1691), "La Metaphysique," 1:120–21; "La Physique, " bk. 8, pt. 2, chs. 1, 24–33; bk. 8, pt. 3.

[18] Antoine Le Grand, *Institutio philosophiae secundum principia de Renati Descartes*, new ed. (London, 1678), trans. Richard Blome, *An Entire Body of Philosophy, According to the Principles of the Famous Renate Des Cartes* (London, 1694; reprint, New York: Johnson, 1972). On Regis and Le Grand, among others, see Hatfield, "Descartes' Naturalism about the Mental," 646–49.

[19] C.C.E. Schmid, *Empirische Psychologie*, 2d edn. (Jena: Cröker, 1796), 189–90.

mind under "moral philosophy," but he nonetheless contended that there are "natural laws" peculiar to mind that can be studied empirically (as in the operation of the senses).[20] In the heyday of Newtonian science, the majority view was that mental phenomena, even if irreducible to material processes, are legitimate natural phenomena.

One should therefore turn a skeptical eye on the easy equating of "naturalism" with "physicalism" in more recent philosophy of mind. Why must the intentional "really be something else" if it is to exist in the natural world? Why shouldn't mentality be taken as a category in nature? We must rethink the easy supposition that these matters were sorted out long ago, so that what remains for us now is simply to discover *how* to reduce the mental to the physical. I am not advocating a return to substance dualism, which has its own problems. However, the early pioneers in the experimental study of the mind, especially of perception, recognized the autonomy and legitimacy of mentalistic notions and rejected the stark alternative of substance dualism or materialistic reductionism. The majority of physiologically oriented psychologists in the late nineteenth-century, including Hermann Helmholtz and Wilhelm Wundt, eschewed physicalism and materialism and effectively adopted an "empirical dualism" that treats the physical and the psychological as two sets of phenomena, each with its own laws. They studied each set of phenomena on its own, and also the relations between them.[21]

The great turn away from this attitude of inclusive naturalism – naturalism that includes mental states within nature – came with the rise of behaviorism in the first third of the twentieth century. All mental phenomena were rendered suspect, as allegedly "ghostly" remainders of a traditional dualism. Without advocating substance dualism, I suggest that it remains open whether mentality might not, after all, be a basic (unreduced) feature of some natural systems, which could remain so even if it were found to be irreducible.

1.3 Machine Psychology

If Descartes and some Cartesians placed the mind in nature, did they do so simply by accepting an immaterial substance as a natural entity, and the laws of mind–body interaction as natural laws? Presumably so. Does this mean that Descartes and the Cartesians placed all the functions of the sensory and the cognitive powers – those portions of Aristotelian psychology that we would now denominate as properly

---

[20] On these eighteenth-century authors, see Gary Hatfield, "Remaking the Science of Mind: Psychology as a Natural Science," in Christopher Fox, Roy Porter, and Robert Wokler (eds.), *Inventing Human Science* (Berkeley: University of California Press, 1995), 184–231.

[21] See Gary Hatfield, "Psychology Old and New," in *Cambridge History of Philosophy, 1870–1945*, 93–106, and *Natural and the Normative*, Ch. 5. On physicalism and mentalism in the first half of the twentieth century, see Hatfield, "Sense-Data and the Mind–Body Problem," in Ralph Schumacher (ed.), *Perception and Reality: From Descartes to the Present* (Berlin: Mentis Verlag, 2004), 305–31.

psychological – on the side of the mind? I think not. Rather, Descartes placed properly psychological functions on both sides of the mind–body divide. He hypothesized that some psychological functions, including forms of learning and memory, could be carried out by mere machines, such as he considered nonhuman animals to be, independent of a mind or any mental properties.[22] Human beings are, in his view, mind–body composites. Even here, he allowed that the human body working alone can perform some psychological functions in a purely mechanistic way; but human beings also possess forms of learning and memory (among other things) that require a mind. In human beings, psychological functions occur autonomously on both sides of the mind–body divide, and also arise (as with sensations) through mind–body interaction.

In the *Treatise on Man* (1664) and *Discourse* (1637), Part Five, Descartes attributed great ability to mindless human bodies and to soulless and therefore mindless animals. Taken together, these abilities allegedly would explain how a mindless machine could adjust its current behavior (i.e., its functionally guided bodily motions) to its current environment in a manner that tends toward sustaining and preserving the machine-body. In these works, Descartes employed the literary trope of a mindless but functioning human body in order to illustrate the capacities that could be ascribed to the body alone. He of course held that human beings essentially have both mind and body, and that the mind leaves the body when the body permanently stops functioning (AT 11:330–31). For human beings, discussions of mindless but functioning machine-bodies describe a counterfactual situation, whereas nonhuman animal bodies are, in his view, mindless machines.

Various of the capacities Descartes attributes to mindless human bodies are psychological.[23] They involve (a) body-preserving behavior that is contingent on current sensory stimulation (AT 11:192–93); (b) changes in the stimulus–response pairings based on the results of previous behavior (11:174, 185); and (c) associative pairings of stimulus events (11:179). In short, he attributed a psychology of instinct, learning, and memory to mindless machines.

Descartes did not himself use the term "psychology". He did, however, discuss the various powers that Aristotelian *De anima* theories attributed to the human soul, and these theories *were* termed "psychology" in his day (though usually by

---

[22] Some recent commentators, including John Cottingham, "Descartes' Treatment of Animals," in J. Cottingham (ed.), *Descartes,* Oxford Readings in Philosophy (Oxford: Oxford University Press, 1998), 225–33, and Stephen Gaukroger, *Descartes: An Intellectual Biography* (Oxford: Oxford University Press, 1995), 278–90, have questioned whether Descartes denied sentience and/or intentionally conceived cognitive operations to nonhuman animals. For the purposes of this chapter, I am simply assuming that Descartes held that animals are mere machines, lacking intentional states. I respond directly to the textual and interpretive arguments of those who disagree in Gary Hatfield, "Animals," in J. Carriero and J. Broughton (eds.), *Companion to Descartes* (Oxford: Blackwell, 2007), 404–25.

[23] A fact recognized by the historian of psychology Raymond Fancher, *Pioneers of Psychology* (New York: Norton, 1979), 21–6.

Protestants).[24] *De anima* psychologies considered the human soul to be the form of the human body and they attributed three to five basic powers to it. In the simple three-fold version Descartes described, these were the vegetative (i.e., vital and reproductive), the sensitive, and the rational powers (AT 1:523; 6:46). Descartes boldly announced that he could explain everything about human and animal life and movement that the Aristotelians had covered but without needing to appeal to vegetative or sensitive souls or powers (AT 6:46, 55–6; 11:200, 202). He sought to fully mechanize the office of the sensory soul for nonhuman animals, and to partially mechanize the sensory power for humans by splitting the psychological functions of that power, such as memory and learning, between mind (a rational or intellectual power) and body.

Let us consider in slightly more detail the functions that Descartes believed the human body could perform without any help from the mind or soul. In the *Discourse*, he summarized portions of the as yet unpublished *Treatise on Man* (initially drafted 1630–33), as well as the lost or unwritten treatise on the soul that was to accompany it. He described the "functions" of a hypothetical "body of a man entirely resembling our own" (AT 6:45–6), composed only of matter lacking the scholastic "forms and qualities" (6:43), and lacking a "rational soul or any other thing to serve as a vegetative or sensitive soul" (6:46). In this soulless human body, powered by a "fire without light," Descartes described various functions "that may occur in us without our thinking of them, and hence without our soul contributing to them" (AT 6:46). Having described the circulation of the blood and the formation of the subtle, but wholly material, "animal spirits," he described what we and (some of) his contemporaries would call psychological functions:

> I also indicated what changes must occur in the brain in order to cause waking, sleep and dreams; how light, sounds, odors, tastes, heat and all the other qualities of external objects can imprint various ideas in the brain through the mediation of the senses; how hunger, thirst, and the other internal passions can also send their ideas there; which part of the brain must be taken to be the common sense, where these ideas are received; which part the memory, which preserves them; and which part the phantasy, which can change these ideas in various ways and compose new ideas from them, and, by the same means, in distributing the animal spirits to the muscles, can make the members of this body move in as many

---

[24] On the term "psychology" in the seventeenth and eighteenth centuries, see Eckart Scheerer, "Psychologie," in Joachim Ritter (ed.), *Historisches Wörterbuch der Philosophie* (Basel: Schwabe, 1971), vol. 7, cols. 1599–1653; Fernando Vidal, "Psychology in the Eighteenth Century," *History of the Human Sciences*, 6 (1993), 89–119; and Hatfield, "Remaking the Science of Mind". The earliest works that adopted the name "psychology" focused on the vital function of the Aristotelian soul in supporting life, in discussing the theology and metaphysics of how the soul is introduced into maternal matter to form a new life; but even those works, along with many subsequent works that did not share this focus, included under the term "psychology" the sensitive and rational powers of the Aristotelian soul. Some historians of psychology deny that any useful comparison can be made between this early "psychology" and the late nineteenth- and early twentieth-century discipline of psychology, apparently treating the term "psychology" homonymously in these contexts: e.g., Roger Smith, "The History of Psychological Categories," *Studies in History and Philosophy of Biological and Biomedical Sciences* 36 (2005), 55–94. I respond to such historiographical worries in Gary Hatfield, "Did Descartes Have a Jamesian Theory of the Emotions?" *Philosophical Psychology* 20 (2007), 413–40.

> different ways as the members of our bodies can move without being guided by the will, and in a manner that is just as appropriate to the objects that are presented to the body's senses, and to the internal passions that are in the body.[25] (AT 6:55).

When this hypothetical machine needs food and has an apple before its eyes, it will eat. When it needs liquid and some is present, it will drink. When a wolf or other dangerous thing approaches, it will retreat (AT 7:229–30). It will, in other words, show behavior appropriate to preserving its survival and meeting its basic needs, which was just the office of the sensitive soul (or power) in Aristotelian psychology (AT 11:192–95).

The passage from the *Treatise on Man* to which this list corresponds made an even stronger claim for the behavior of this mindless machine. Having mentioned mechanisms underlying sensory functions, imagination, memory, and the appetites and passions, he claimed that these mechanisms were sufficient to explain "the external movements of all the limbs, which follow so appropriately both from the action of objects presenting themselves to the senses and from the passions and the impressions found in the memory that they imitate as perfectly as is possible the external movements of a real human being" (AT 11:202) – and do so in clockwork fashion, without any sensitive soul, or anything not found in inanimate bodies. Indeed, earlier in the *Treatise* he had boasted that "the effect of the [corporeal] memory which seems to me here the most worthy of being considered is that, without there being any soul in this machine, it can naturally be disposed to imitate all the movements that real human beings, or else other similar machines, will make when the soul is present" (AT 11:185). Here he claims that he can explain, in a purely mechanical way, all human behavior. Corporeal memory plays the starring role. By influencing the control of behavior in response to previous sensory inputs, corporeal memory adjusts behavioral responses so as to adapt them to changing circumstances. In this way Descartes would explain how animals, such as dogs and birds, can be trained to respond to various sensory cues (AT 4:574–75, 11:370).

In the *Discourse*, Descartes mitigated the strong claims he had made for mindless human machines in the *Treatise*. Immediately following the long passage quoted above, he famously provided two "very certain" means by which real human beings having minds or souls can be distinguished from mere automata. These well-known criteria are (1) the ability to use words to express thoughts, and (2) general reasoning ability that can be applied to novel situations. The first might be called "conceptual language" to distinguish it from the mimicry of parrots. The second might be called "general intelligence". To these we would need to add Descartes' own third and fourth distinctive aspects of mentality, although not asserted in this context:

---

[25] Descartes' use of the term "idea" in this passage might suggest that he has mentalized this machine; however, in the 1630s he utilized the notion of a *corporeal idea*, a bodily state that served to control movement (in a sensorimotor feedback loop), and which, if the body were ensouled, might cause a properly mental idea (AT 11:176–77); in the Objections and Replies to the *Meditations* (AT 7:160–61), he explicitly restricted this sense of the term "idea" to instances in which the mind is actually affected by the brain state, and so he now refused to term the brain pattern by itself an "idea" (as he had done in the *Discourse* passage here quoted). The "phantasy" is a brain organ, also known as the corporeal imagination (AT 7:160–61, 10:414).

conscious experience and free will, which he believed depend on an immaterial mind for their existence.

For now, I'm more interested in what mindless machines *can do* (according to Descartes) than what they cannot. Descartes' views may be summarized by comparing the psychological functions he ascribed to his mindless machines with the classification of sensory and rational functions in Aristotelian psychology. In effect, Descartes redistributed the Aristotelian psychological and mental functions across his mind–body divide. Retaining the basic functional language of Aristotelian psychology, he denominated sense, memory, imagination, and reason (or the intellect) as distinct capacities.[26] He, as did the Aristotelians, further divided sense into external and internal, and placed the appetites as internal senses. He also recognized the category of passions or emotions, although, unlike Aristotelians such as Aquinas,[27] he did not treat the passions as appetites but set them up as a distinct category (AT 11:346–48). He then distributed these psychological functions according to his new mechanist and dualist framework, in the manner displayed in the following Table.

Cartesian Psychological Functions Across the Mind–Body Divide

Animals, human body alone (purely physiological)	Human mind–body complex	Mind alone
Sensory-motor instincts	Sensory & motor awareness (bodily effects on mind)	Pure intellect (clear & distinct ideas)
		Judgment & reasoning (intellect & will together)
Corporeal imagination or "phantasy"	Experienced mental images (brain images affect mind & mind can induce images)	Intellectual grasp of spatial extension (non-imagistic innate ideas)
Memory-based sensory-motor responses that influence bodily motions (learning)	Acquired mental habits relating to senses and bodily motion	Intellectual memory
Bodily "appetites" (internal sense)	Feelings of appetite (int. sense: hunger, thirst)	Acts of free will
		Purely mental appetites
Bodily "passion"	Mental ("proper") passion, i.e. bodily caused feeling	Purely mental emotions

---

[26] Sensory functions may be either purely bodily (in mindless machines, or in human behavior unguided by mind), or involve mind–body interaction. Descartes granted two powers to the mind: intellect and will (Pr I.32). Mental states of sense and imagination are effects of the body on the intellectual power (AT 7:78–9). In imagination, the mind can induce brain images or can experience images arising from the brain (AT 11:344–45, 348–49). Corporeally based memory differs from intellectual memory (AT 3:48, 84). Pure intellect operates independently of bodily influence and material images (AT 7:72–3). The will may respond either to sensory materials or to purely intellectual representations (Pr I.32–4, AT 7:58–9, 65, 72–3, 76). On corporeal and mental appetites and passions, see Hatfield, "*Passions of the Soul* and Descartes' Machine Psychology," 3, 14–6.

[27] Thomas Aquinas, *Summa Theologica*, trans. Dominican Fathers (London: Oates & Washbourne, 1922), Pt. I, qq. 80–1. On Aquinas' theory, see Simo Knuuttila, *Emotions in Ancient and Medieval Philosophy* (Oxford: Clarendon Press, 2004), 239–55.

This Table makes plain that, for Descartes, the psychological and the mental are not coextensive. He believes that some functions that he (and we) would classify as properly psychological can be fully mechanized without invoking mentalistic notions. Other properly psychological functions can be partially mechanized, and therefore mentalistic notions are only partially implicated in their explanation. Still others, the pure intellect and the will, are purely mental. According to him, then, construing a function as psychological does not make it mental. Careful study of this Table forces us to acknowledge that Descartes posited a relation between the psychological and the mental that differs from many recent conceptions.

But can Descartes make it work? That is, leaving aside the physiological falsity of his particular mechanisms, could purely mechanical structures plausibly account for the sensory and motor capacities of animals, and for nonmentalistic human behavior? To assess the plausibility of his position, we need to consider his conception of the linkage between external objects and organismic responses.

Throughout his physiological writings, from the *Treatise* to the *Passions*, Descartes used the language of causal correspondence to describe the relations among external objects and their effects on the sense organs, nerves, and brain. The shape and the color of an object variously affect light, which variously affects the eyes, nerves, and brain. Differences in the object are reflected in differences in brain state. The same object under the same internal and external conditions will produce the same brain state on various occasions (AT 11:174–95; 11:338–39).

The idiom of causal correspondence was also popular among behavioristically inclined psychologists and philosophers of the twentieth century. On this view, intentional powers of representation need not be ascribed to brain states in order to account for the effects of those brain states in mediating between external stimuli and behavior. In order to explain situationally appropriate behavioral responses, we need only to assume the presence of three factors: first, a correspondence between various external states (e.g., presence of food, or presence of a predator) and internal sensory states (in the brain); second, the existence of other internal states (e.g., hunger, mechanically defined with reference to the stomach or blood) that modify the response to a given sensory state; third, the capacity to change the alignments between sensory states and behavioral responses (learning and memory, mechanically conceived).[28] In this way, the dog (or human) might learn to associate the presence of food with the presence of a nonedible stimulus, such as the ringing of a bell. If "learning" can mean relatively stable changes in patterns of behavior as a result of stimulus history, learning might in principle be accounted for by changes in the plumbing (and/or wiring) that mediates between sensory input and motor response.

Now that behaviorist psychology has waned, the question of how far a behavioristic analysis of learning might successfully extend is seldom asked. Human

---

[28] These factors align with the physiologically oriented behaviorism of J.B. Watson, *Psychology from the Standpoint of a Behaviorist* (Philadelphia: Lippincott, 1919), and Clark L. Hull, *Principles of Behavior: An Introduction to Behavior Theory* (New York: Appleton-Century, 1943), as opposed to the "hollow organism" behaviorism of B.F. Skinner. See also Hatfield, "Behaviorism and Psychology".

mental life contains phenomena that require the application of mentalistic concepts for their description. Conscious experience is among the most obvious: it would appear to be a stable feature of (at least) human mental life (see Section 1.4.), and yet it was left out of the standard behaviorist catalogue of natural phenomena. More generally, a telling argument against twentieth-century behaviorist explanations of perception was that such explanations could not make sense of the notion of perceptual error, which entails a notion of *misrepresentation*. If an animal can *misrepresent* the presence of food, it must be able to *represent* food in the first place, and so it must possess intentionality. Misrepresentation demands the concept of representation or intentionality.[29]

In order to review the plausibility of Descartes' machine psychology on its own terms, we need not consider arguments for or against a role for intentionality in scientific or philosophical accounts of sense perception. For, as it happens, Descartes did not believe that the notion of "error" was appropriately applied to his mindless machines. As he said in the Sixth Replies, "no falsity can occur" in the purely physiological processes he there labels as the "first grade of sense" (AT 7:438). Error proper depends on the concept of judgment, and arises with minds (for him, as for many others).

Nonetheless, to retain plausibility, Descartes' machine psychology must accommodate maladaptive responses by the sensory-motor systems of his mindless machines. And indeed, we normally think of machines as things that can be broken, as things that may function properly or improperly. The machine metaphor thus might allow us to say that a sheep that doesn't run from the wolf is in some way broken (relative to a notion of proper physiological functioning). But being broken is not the same thing as being in error. Hence, if machine psychology is given the resources of functional description (something to which Descartes helped himself, whether he was entitled to do so or not),[30] then such a psychology can descriptively accommodate dysfunctional responses in sensory or other mechanisms without recourse to the mentalistic notion of misrepresentation. A nonmentalistic psychology that includes a concept of proper functioning might be adequate for at least some nonhuman animals, and might even cover some aspects of human psychology.

For analyzing Descartes' machine psychology, behavioristic descriptions offer a useful conceptual tool. While I would agree with those, including Descartes, who do or would hold that behaviorism is not an adequate general methodology or philosophy of psychology, it may well be that behavioristic conceptions retain their usefulness in some domains. Behaviorist conditioning works, and it may in some cases work simply as associative conditioning, rather than by means of cognitive mechanisms that mirror the results of purely associative mechanisms. In that case, the extent to which an updated machine psychology might apply to some human and

---

[29] See Fred Dretske, "Misrepresentation," in R. Bogdan (ed.), *Belief* (Oxford: Oxford University Press, 1986).

[30] On the notion of function and malfunction in Descartes' physiology, see Gary Hatfield, "Descartes' Physiology and Its Relation to His Psychology," in John Cottingham (ed.), *Cambridge Companion to Descartes* (Cambridge: Cambridge University Press, 1992), 335–70, and "Animals".

animal behavior remains open. The history of such a machine psychology would pass from Descartes, through T. H. Huxley,[31] to Watson and the behaviorists proper. That history, too, forms part of the history of theories of mind.

1.4 On the Legitimacy of Introspection

According to a standard account, Hume and Kant discredited introspection as a method for studying the self and its activities, and the behaviorists discredited introspection as a psychological method for studying experiential contents. Hume and Kant successfully countered the (allegedly) Cartesian idea that the mind or its operations are fully transparent to introspection: when we look within, we do not find a "self" as simple substance open to our gaze, nor are the operations of the mind fully accessible. The behaviorists and their philosophical allies subsequently argued that any sort of introspection was unscientific, or conceptually impossible, or akin to believing in ghosts.[32]

Among philosophers there has been a recent revival of interest in self-knowledge of some mental states – especially opinions and convictions. Broadly construed, such first-person knowledge might be called a kind of introspection, though its advocates eschew a "perceptual model" of reporting on the character of one's current experience.[33] Indeed, the notion of introspective awareness of phenomenal aspects of perceptual states remains deeply suspect, for reasons I shall discuss.

The force of these various criticisms, and any assessment of the extent to which introspection has been discredited, are highly sensitive to what one takes the objectives of introspection to be. Thus, if introspection was supposed to reveal the soul as a simple substance, it clearly fails. Similarly, if introspection was alleged to provide incorrigible evidence about the nature of psychological states and the processes underlying them, it fails again. But if introspection is intended simply as a report on or response to the character of present experience, then, for at least some aspects of experience, introspection becomes plausible: for reporting whether we

---

[31] T.H. Huxley, "On the Hypothesis that Animals Are Automata and Its History," *Science and Culture* (New York: Appleton, 1884), 206–54.

[32] For an instance of the view that Descartes thought introspection could yield "transparent" insight into the mind as simple substance, see Owen Flanagan, *The Science of the Mind*, 2nd edn. (Cambridge: The MIT Press, 1991), 16; on the more general claims of "Cartesian transparency," 194–200. On the Humean (and Kantian) critiques, see Sydney Shoemaker, *The First-Person Perspective and Other Essays* (Cambridge: Cambridge University Press, 1996), Ch. 1. For a standard account of the demise of introspection, see William Lyons, *Disappearance of Introspection* (Cambridge: The MIT Press, 1986), Ch. 1.

[33] Richard Moran, *Authority and Estrangement: An Essay on Self-Knowledge* (Princeton: Princeton University Press, 2001). Shoemaker, *First-Person Perspective*, Ch. 1 and Part IV, also derides a "perceptual" model of introspection. The journal *Philosophical Topics* 28:2 (2000), is devoted to the topic of introspection, and contains several articles on self-knowledge. In the same issue, David Rosenthal, "Introspection and Self-Interpretation," 201–33, defends a broader than usual notion of introspection (closer to that defended herein).

are currently experiencing an appearance of red, or are having the appearance of something moving in the field of vision, or are having the smell of an apple pie at close range, and so on. The question of the objective of introspection, of what it is supposed to reveal, becomes paramount in assessing claims about its legitimacy.

Here again, attention to the history of theories of mind can help sort out the present-day landscape in philosophy. Consider the philosophical criticism of the introspective search for the soul as simple substance. This criticism loses its sting when we realize that Descartes did not claim to find the soul, let alone its nature as a substance (simple or otherwise), via introspection. He did claim to become aware of his "thoughts" – sensations, images, judgments, volitions – via introspection (AT 7:28–9), thereby becoming aware of *instances* of the attribute of thought. That is not the same as being directly aware of the mind as an immaterial substance. Descartes' metaphysical conclusion that the mind is a simple, immaterial substance does not arise simply from introspection, but from his (alleged) intellectual insight into thought and extension as the essential attributes of substances (AT 7:175–76; 8A:25). That is, in concluding that mind is distinct from body, Descartes did not simply compare his awareness of his own thoughts with his perceptions of extended things. Rather, he argued from his clear and distinct intellectual comprehension of thought and extension as attributes of complete beings to the conclusion that mind and body are mutually distinct kinds of substance (AT 7:78, 119).

We may reject some of the premises of Descartes' argument and its conclusion, and we might therefore come to believe that Descartes' alleged intellectual insights did not in fact go beyond mere introspection. But that is different from showing that Descartes framed his arguments as introspective reports of a direct inner inspection of the soul in its immateriality – something he never claimed to achieve (nor did any other major philosopher). In this instance, we have found that a "standard" philosophical objection to introspection was misdirected, so that if we reject Descartes' argument, we must do so on others grounds (e.g., by denying that the pure intellect can directly comprehend the essences of substances through a kind of intellectual insight).

Another standard objection against allegedly "Cartesian" introspection rightly observes that the mind's processes and operations are not fully "transparent to thought". This objection presupposes that Descartes or another major philosopher held that the mind's processes were thusly "transparent". In fact, Descartes held no such thing. He did not think that we can discover the mind's most basic operations through simple introspection. According to him, the optimum mental states are "clear and distinct perceptions" by the intellect. Also according to him, most people go through their entire lives without ever having such a perception (AT 7:477). It takes work – work guided by a hard-won and correct theoretical perspective – to uncover such intellectual perceptions. Descartes aims to help his readers to achieve those sorts of perceptions in the *Meditations*. Far from simply asking his readers to introspect, he asks them to undergo certain cognitive exercises in order to reveal what has been hidden to them: that they are capable of clear and distinct perceptions of the essence of mind, matter, and God (AT 7:9–10). If he thought the mind were "transparent," there would be no need for such work. More generally, Descartes

also held that many ordinary operations of the mind, such as the rapid or habitual judgments that underlie sense perception, are unnoticed by perceivers (AT 7:438). He posits such judgments on theoretical grounds, not on the basis of introspective evidence of their occurrence.[34]

If it were the case that philosophers such as Descartes had claimed to discern that the soul is a simple, immaterial substance simply by looking within, or if philosophers had truly claimed that the mind's operations are fully "open" to introspection, then objections would be in order. But if no philosopher claims that sort of introspective insight, then many would-be critics need to go back and accurately characterize what role introspection is playing before tendering their criticisms. In Descartes' case, he claimed to be aware of the experiential character of various "thoughts," a term he used in a wide sense to include a variety of states of mind. He claimed to be able to know, introspectively, that he had occurrent thoughts or experiences of various kinds, including sense perceptions, images of imagination, and acts of will. This sort of introspection, of phenomenal aspects of current experience, became important in subsequent psychology.

The notion that thoughts or perceptions have an introspectable "experiential character" is under attack today. As in the previous case, this attack depends on a certain construal of what introspection is supposed to show. According to its critics, introspection is supposed to be a special "inner" awareness, which is directed at an object that is distinct from external (material or physical) objects. And yet, the criticism goes, when observers are asked to introspect their experience of a red ball, they find only the awareness of a red ball; they do not find another object, separate from the experienced red ball, within their introspective gaze. In other words, our experience of the ball seems to be of something "out there," not "in the head". Hence, the critic contends, there is nothing special for introspection to report; it tells us nothing that differs from ordinary descriptions of what we see when we look at a red ball.

Georges Rey ably summarizes this sort of criticism: "There are a good number of problems raised by introspectivist proposals. In the first place, as a number of writers (e.g. Harman 1990, Dretske 1995) have stressed, a great deal of what passes for introspection of one's 'inner' experience consists of reports about how the *outer* world seems: we don't so much report on the features of the 'inner movie' as upon what that movie *represents* (e.g. that *barns* seem red, *the sky* a dome)".[35]

This observation about the apparently external character of the objects of perception has been labeled the "transparency" of perception, where that term is now used to connote the (purported) fact that perception has no character of its own. Perceptual states are supposed to put us "transparently" in touch with the external

---

[34] See Gary Hatfield and William Epstein, "The Sensory Core and the Medieval Foundations of Early Modern Perceptual Theory," *Isis* 70 (1979), 363–84. On the *Meditations* as cognitive exercises, see Hatfield, *Descartes and the Meditations* (London: Routledge, 2003), Ch. 2.

[35] Georges Rey, *Contemporary Philosophy of Mind*, 136–37. The references are to Gilbert Harmon, "The Intrinsic Quality of Experience," in J.E. Tomberlin (ed.), *Philosophical Perspectives 4: Action Theory and Philosophy of Mind* (Atascadero: Ridgeview, 1990), 31–52, and Dretske, *Naturalizing the Mind* (see 57–62).

world. Because the ball seems to be "out there," our perception of it allegedly does not involve any subjective experiential component. The ball is just there, and we see it. We "see through" our perceptual states to the world. There is nothing to introspect beyond the report of world-states, described in the language of object-properties. In reporting on the presence of red in the visual field, we are simply stating the presence of a certain physical property. There is no subjectively phenomenal aspect for us to be aware of.[36]

This (ever more popular) "transparency" argument against introspection improperly runs together a number of issues, both historical and philosophical. In my view, the specific philosophical issues do not bear either way on the legitimacy of introspection. However, the historical factors help to explain how philosophers could mistakenly believe that the philosophical issues did bear on that question.

Let us first consider a thesis about the metaphysics of color qualities that is embedded in this argument. In reporting the color of the red ball, the question of whether the content of one's report is exhausted by the mind-independent properties of world-states depends on one's metaphysics of color. According to some theorists, including Michael Tye and Fred Dretske, in experiencing color we are simply instantiating the informational content that a certain physical property is present.[37] These philosophers believe that there is no subjective feature to this experience: the objective content, the physical property of the object, exhausts the content of the perception.

Among the alternatives to this metaphysical position on color qualities is my own favorite position, which is called a "relational" or "dispositional" theory. On this view, the experience of red is subjectively conditioned by our perceptual apparatus. To see a colored thing is to have an experience of a certain kind, the character of which varies depending on what species we belong to (human, feline, etc.) and whether our perceptual apparatus is "normal" for our species. The character of that experience is essentially an object of introspection: it cannot be deduced simply from the mind-independent properties of the surface of the ball.[38]

The question of which metaphysics of color is correct – whether in seeing a color we are directly aware of a physical property, or whether we are made aware of the ball through a subject-dependent experiential color-feature – cannot be decided

---

[36] Similar arguments about perceptual transparency are offered by Michael Tye, *Consciousness, Color, and Content* (Cambridge: The MIT Press, 2000), 45–54, and John Campbell, *Reference and Consciousness* (Oxford: Clarendon Press, 2002), Chs. 6–7. By contrast, Tim Crane, "Introspection, Intentionality, and the Transparency of Experience," *Philosophical Topics* 28 (2000), 49–67, while agreeing with the "intentionalist" theory of perceptual qualities of Dretske and Tye, denies that the theory implies "transparency," and questions whether phenomenological observation could settle this point about the metaphysics of perceptual content. Sydney Shoemaker, "Introspection and Phenomenal Character," *Philosophical Topics* 28 (2000), 247–73, endorses intentionalism (with Dretske, Tye, and Crane), but affirms introspectable qualia.

[37] Tye, *Consciousness, Color, and Content*, 45–54; Dretske, *Naturalizing the Mind*, 88–93.

[38] Gary Hatfield, "Objectivity and Subjectivity Revisited: Color as a Psychobiological Property," in Rainer Mausfeld and Dieter Heyer (eds.), *Colour Vision: From Light to Object* (Oxford: Oxford University Press, 2003), 187–202.

through an analysis of the phenomena. That is, even if we were all to agree that it *seems* (phenomenally) that color is a simple mind-independent property "out there," that would not settle the metaphysical question, any more than the fact that the Earth *seems* stable to us decides whether the Earth rotates around its axis daily or the Sun revolves around the Earth daily. How things *prima facie* seem to us is decisive neither in physics nor in metaphysics. Hence, I happily concede the experiential point, that color seems to be "out there," not "in here," and then I go on to contend that this concession neither decides the metaphysics of color qualities nor bears ill tidings for introspection.

How, then, could so many critics believe that the fact that objects and colors appear "out there" bodes ill for introspection? I think that *phenomenal externality* (the red is "out there") has been confused with *property transparency* (we are directly aware of the physical property in an object that causes us to see the object as colored).[39] The correct phenomenal point, that the ball appears "out there" (not "in here") has been confused with the contentious claim that, in seeing the ball, we do so by directly apprehending its mind-independent features. But further, and more importantly, phenomenal externality could appear to bear on the legitimacy of introspection only given certain conceptions of the experiential nature and objectives of introspection. In this case, the critics have interpreted the notion that introspection is a kind of "looking inward" by supposing that its objects should seem phenomenally to be "in the head" rather than "in the world". The critics of introspection have confused a metaphysical claim advanced by some of their adversaries on color ontology – that color is "inner" in the sense of being subject-dependent – with the experiential claim that color experience should *seem* to be "inner" rather than "outer". Those who hold that color is a subject-dependent quality (e.g., the relationalists/dispositionalists) do not thereby commit themselves to the view that color should seem to appear "inside" the perceiver. Rather, these philosophers commit themselves to the view that the way colored things look depends essentially on characteristics specific to the human species (and perhaps its close relatives) that determine the way the perceiver's perceptual system presents colored objects in perceptual experience.[40]

All introspectivists do not adopt a particular view (or any view at all) about the metaphysics of color, and so we may set that issue to one side for now. Instead, we

---

[39] "Property transparency" may be further distinguished from "metaphysical transparency," the view that the metaphysics of sensory qualities is manifest to us in virtue of our being perceivers, or by simple reflection on the phenomenology of sensory experience. See also Hatfield, "The Reality of Qualia," *Erkenntnis* 66 (2007), 133–68.

[40] Dispositionalists are often accused of believing that we see sense data, rather than objects; this accusation may lead to the further criticisms that sense data, as distinct entities, should appear phenomenally to be somewhere other than where the object is seen and also that they are unneeded and epistemologically undesirable theoretical posits. But a qualia realist need not be a sense-data theorist, and need not hold that color qualia are "what is seen" or that they are phenomenally distinct from objects; rather, qualia realists can treat qualia as aspects of experience that help explain *how things are seen*. See Hatfield, "Sense-Data and the Mind-Body Problem" and "Reality of Qualia".

must ask why their critics would believe that introspectivists should generally be committed to the experiential claim that the objects of introspection are "in here".

Attention to the history of theories of mind sheds light on how this mistaken attribution occurred. It turns out that early advocates of introspection, including the psychologist and philosopher Wilhelm Wundt, used the terms "inner" and "outer" to denote, respectively, the object of introspection, and objects as observed in the physical sciences. However – and this is crucial – Wundt did not intend his use of "inner" to carry the *experiential* implication that the objects of introspection should seem to be "in here". Rather, he used the two terms to denote two conceptual standpoints from which *the same* external objects might be observed. Here (in translation) is an example of Wundt's language:

> The ideas of which psychology seeks to investigate the attributes, are identical with those upon which natural science is based. [...] It follows, then, that the expressions outer and inner experience do not indicate different objects, but *different points of view* from which we take up the consideration and scientific treatment of a unitary experience.[41]

The phrase "point of view" might suggest that the objects of introspection should be looked for by directing the gaze elsewhere than toward the world. But that is not at all what Wundt meant. "Point of view" denotes a theoretical attitude that is brought to experience.

Wundt's notion of introspection may be illustrated through experiments in visual perception, the area of his own first work in psychology.[42] In studies of visual perception, experimenters vary certain stimulus values and then ask observers to match one stimulus to another, or to report certain features of the stimulus. The oft-repeated complaint against introspection, that the act of introspecting would intrude on the experience and change it, doesn't fit this case. Subjects may be offered some stimuli to look at – say, two color patches, or two lines presented perpendicular to one another – for several seconds. They would be asked to determine whether the two stimuli look the same (or not). From Wundt's perspective, what makes these observations cases of introspection about perception is that the experimenter is interested in how the lines *appear* to the observer, in the lines as features of the observer's experience. The experimenter is not seeking to determine the physical facts about the lines by asking the subjects to serve as lab assistants (she is not asking the subjects to make reliable physical measurements). Nor is the experimenter seeking to discover what the subject *believes*, on the basis of his observations, the real physical properties of the stimulus might be: whether the subject believes that the colors or lines *really are* physically the same. The subject might be able quite easily to determine (because he is familiar with line-illusions) that one line is actually shorter than the other; but the experimenter would nevertheless ask him to report, by recourse to the looks of

---

[41] Wilhelm Wundt, *Outlines of Psychology*, trans. C. H. Judd, 2nd edn. (Leipzig: Engelmann, 1902), 2–3.

[42] Wilhelm Wundt, *Beiträge zur Theorie der Sinneswahrnehmung* (Leipzig: Winter, 1862). Examples of the types of observations described in this paragraph may be found in Wundt, *Outlines*, 72–7, 136.

the lines alone, whether they appear to be the same or different in length. A different experiment might be done to test subjects' inferential beliefs about the lines (based on their knowledge of illusions, or other cognitive factors); but that would be an experiment about inferential belief, not about perceptual experience *per se*.

Consider a more ordinary example. Suppose that you are painting a table red. You buy the paint, watch the salesperson mix the pigment, and watch the can being thoroughly shaken in the usual mechanical device for shaking paint cans. You are assured that the pigment has been homogeneously distributed throughout the liquid paint. Now you paint the table, covering the old surface completely by using two coats. After the paint dries, someone asks you how the table looks. You say, "red". If the person asks whether it looks to be the same red across the entire surface, you will say, without hesitation and truly (assuming a good paint job), "yes".

But now suppose that someone asks you to attend to your experience more closely, to report on its character more finely. She asks you whether the red surface "looks the same" all the way across. Now you notice that the surface looks different at different places: a highlight in one place reflects the windows above the table as distinguishable trapezoids; the light also falls unevenly on the surface, making it appear darker in one place, lighter in another. What you are noticing, what you are reporting, are aspects of your experience. No doubt these aspects have physical causes. Yet, you can report on them in entire ignorance of those causes simply by attending to subtle variations in how the table looks.

The important point here is that you can make your reports on these subtle looks of the table independently of your attitudes toward or beliefs about the metaphysics of experience or of introspection. The grain of your experience changes when you are asked to attend more carefully to the way the surface of the table looks. Now you are attending to how the table looks in ways that you would not attend to if you were primarily concerned with whether the same paint was used to cover the entire surface. You might determine *that* physical fact quite easily without closely attending to highlights or to shadows, or by discounting their appearance.

Introspection of the sort that involves describing one's experience is rampant in the literature of perceptual psychology. A well-known instance occurs in the figures that the Gestalt psychologists offer as demonstrations. The Gestaltists asked their readers to pay attention to figure-ground reversals and to various phenomena of grouping that arise when they look at line drawings.[43] The drawings remain physically the same, yet the character of the observer's experience changes. Subjects report these changes quite easily, and, in so doing, they are reporting on their experience. It remains true that they are reporting on their experience no matter what the correct metaphysics of experience might be. Plainly stated, a role for introspective reports exists, independently of whether property transparency is true. Property transparency cannot be established by experiential description of the

---

[43] E.g., Kurt Koffka, *Principles of Gestalt Psychology* (New York: Harcourt, Brace, 1935), 83, 159, 164, 195.

objects of perception as seeming to be "out there". That experiential point can be accepted, while the metaphysical question remains open.

My conjecture is that the recent critics of introspection simply misunderstood the intent of authors such as Wundt who spoke of "inner" and "outer". These critics interpreted Wundt (or others) as saying that the objects of introspection should seem to be "in here". They then took this construal to provide the test case for deciding whether anything was there to introspect.[44] This conclusion gained support from their confounding property transparency with experiential externality. Mistakenly, these critics believed that their metaphysical claims about the content of phenomenal experience – that its content is exhausted by the representation of mind-independent physical properties – could be supported by the experiential observation that the objects of experience seem to be outside the body and hence outside the mind. Once this confusion is cleared up, it should be apparent that a belief in the efficacy of introspection can be held independently of various metaphysical positions about color qualities or the content of perception.

1.5 Conclusions

The history of theories of mind is interesting in its own right as a branch of the history of philosophy and the history of the psychological sciences. Work in this field also can remedy confusions that lead to mistakes or misdirected arguments in contemporary philosophy of mind. We have seen instances of this with respect to assumptions about "standard views" of the place of mind in nature, the division between the mental and the psychological in early modern machine psychology, and the objectives of introspection. Some of the confusions resulted from the unacknowledged but continuing influence of behaviorist attitudes that became entrenched in philosophy during the middle decades of the twentieth century.

At the same time, behavioristic concepts have provided useful analytical tools in our discussion of machine psychology. There is a need for explicit discussion and reconsideration of the place of behaviorism in the history and philosophy of theories of mind and in the attendant historiography. Any philosophical reconsideration should examine the plausibility of behaviorist conceptions of learning and memory with respect to both human and animal psychology. Even if, as I believe,

---

[44] The earlier critics of introspection, including Watson, *Psychology from the Standpoint of a Behaviorist*, charged that introspection was scientifically unreliable. They pointed to the controversies over such questions as the number of "levels of attention" and the number of basic sensations (overall, or within a given domain, such as color). These criticisms were apt. However, the Gestalt psychologists, who also relied on reports of experience, offered some of these same criticisms. The problems that the behaviorists noted were largely restricted to the "analytic introspection" favored at the turn of the twentieth century by E. B. Titchener and others, on which see Hatfield, "Introspective Evidence in Psychology," in Peter Achinstein (ed.), *Scientific Evidence: Philosophical Theories and Applications* (Baltimore: Johns Hopkins University Press, 2005), 259–86.

behaviorist conceptions are inadequate as a general account of human and animal capacities, they may well be adequate and useful in some domains.

In the cases considered herein, the history of theories of mind and of the use of psychological concepts in philosophy was valuable in clarifying contemporary debate by contrasting received views about historical positions with the views that the theorists in question actually held. Such work is of contemporary relevance because statements about the positions of major figures such as Descartes or Wundt are often used to set a standard framework within which debates then occur. If such frameworks are inaccurate, our confidence in using them to frame debates is reduced. If the positions constituting the framework are misstated so as to make them easy targets for refutation, we should be skeptical of whatever support is alleged to accrue to a position through its refutation of such straw men.

In the end, contemporary theorists are left with work they themselves must do in deciding whether unreduced intentionality can be included within nature, whether behavioristic analyses of learning have an ongoing role to play, and whether introspection is legitimate. By directing contemporary theorists away from apparently easy, but ill-gotten, refutations of their own least favorite positions, historical reflection can help all of us to see that these topics remain open, and it may even uncover conceptual resources that have been neglected (or misused) of late. On the assumption that the cases studied herein are only a few among many that might be investigated, historians of psychology and of theories of mind have plenty of work to do, with plenty of benefits for contemporary theory.

Chapter 2
Philosophical Psychology in 1500: Erfurt, Padua and Bologna

Pekka Kärkkäinen and Henrik Lagerlund

Abstract The chapter gives a general description of philosophical psychology as it was practiced and taught in the sixteenth century at three of the most important universities of the time, the universities of Erfurt, Padua, and Bologna. Contrary to received notions of the Renaissance it argues that the sixteenth-century philosophical psychology was tightly bound to the Aristotelian tradition. At the University of Erfurt, philosophical psychology was developed with strong adherence to the basic doctrines of Buridanian *via moderna*, as it had been taught for over a century. The Buridanian approach dominated especially discussions on the metaphysical nature of the human soul and disputes about universal realism versus nominalism. The situation was somewhat different at the universities of Bologna and Padua. The connections between these two universities were close, and they can be seen as developing one and the same Aristotelian tradition. Although the works produced were rather eclectic in nature, they shared research topics as well as conceptual and methodological frameworks which contributed to the unity of the school. In Bologna and Padua, Averroës had a central position as an authority cited and criticized; and philosophical questions concerning the immortality of the soul and the nature of the intellectual species attracted continuous interest. The development of philosophical psychology was also influenced by the special organizational situation of these universities: theology had a relatively unimportant position, and medicine instead had continuous impact on teaching.

Keywords *De anima* · doctrinal history · immortality of the soul · nominalism · Usingen · Averroës · Pomponazzi

The question we ask ourselves in this chapter is simply: What was philosophical psychology like around 1500? Perhaps the more appropriate question is: What kind of philosophical psychology was taught? We would thus like to provide a general description of philosophical psychology in 1500. We have chosen this date since it falls between two great periods in the history of Western philosophy, namely, the flourishing intellectual activity of the thirteenth and fourteenth centuries and the

P. Kärkkäinen
Department of Systematic Theology, University of Helsinki

so-called birth of modern philosophy in the seventeenth century. Our main hypothesis is that a study of theories of the mind from the time around 1500 will support the claim that there is an overreaching continuity between these periods.

This is an ambitious project and not something that could be easily presented in one chapter. The present study should instead be seen as a starting point of a larger project. To make the project easier to handle, we focus on three of the most important universities of the time, namely, Erfurt, Padua and, to a lesser extent, Bologna. In this way, we hope to present a representative picture of the subject – if only at this point a rather sketchy one.

We will proceed by presenting how and what was taught in philosophical psychology at these universities, and we will also discuss some of the more important debates and problems that were explored within the most significant works of the time. The nature of this project is, thus, more historical than philosophical. It aims, after all, at an accurate historical picture of philosophical psychology around 1500, but it should be emphasized that without paying particular attention to the details of the philosophical theories discussed at the time such a picture will be hopelessly inaccurate.

In this discussion, we frequently use the words "philosophical psychology". What we mean by these words will be clear, we think, but it still seems appropriate to mention that they denote the part of natural philosophy which in 1500 dealt with the nature of the soul and its operations – basically what Aristotle deals with in his treatise *De anima*. It needs hardly to be stressed that the problems involved in presenting a workable theory of the human soul were different in 1500 than at the time Aristotle gave his lectures on the soul, but as a rough guide to what we mean by the words "philosophical psychology" the above note is sufficient.

2.1 Philosophical Psychology at the University of Erfurt

During the fifteenth century the University of Erfurt was one of the great intellectual centres of Germany. Its Faculty of Law was widely respected, and its theologians gained prominent positions in the Church.[1] Above all Erfurt was, along with the University of Vienna, a university where philosophy was taught exclusively according to the principles of the *via moderna*,[2] and it remained practically unaffected by

---

[1] See Robert Gramsch, *Erfurter Juristen im Spätmittelalter* (Leiden: Brill, 2003), 69–71; Erich Kleineidam, *Universitas studii Erffordensis II* (Erfurt: Benno, 1992), 3–4.

[2] Recent research emphasizes that different schools of thought were distinguished in the fifteenth century by their differing methods of reading Aristotle, which were articulated by an appeal to authoritative expositors of each school. Buridan's commentaries were exemplary for the *via moderna.* An important characteristic of the method of this school was to attribute the definitions to terms and not to things, and, because of this, its adherents were frequently called terminists or nominalists. This applies also to the way the Erfurtian philosophers understood themselves. See Wolfgang Urban, "Die '*via moderna*' an der Universität Erfurt am Vorabend der Reformation," in Heiko A. Oberman (ed.), *Gregor von Rimini, Werk und Wirkung bis zur Reformation* (Berlin: De Gruyter, 1981), 327, n. 54. On late medieval schools of thought in general, see Maarten J.F.M. Hoenen, "*Via*

the miseries of the so-called "Wegestreit". Erfurt had retained its traditional way of teaching that followed for the most part the mainstream of fourteenth century teaching of philosophy, including that of John Buridan, William Ockham, and their followers.[3] However, during the early years of the sixteenth century, the traditional form of teaching was challenged first by humanism and later by the Reformation. At first, it seemed that a peaceful co-existence of the old and new ideas would be possible, but a course of events resulted in the decline of the whole university by the early 1520s.[4]

2.1.1 Writings on Natural Philosophy in Late Medieval Erfurt

Philosophical psychology was taught as a part of natural philosophy. In 1482, printer Paulus Wider de Hornbach distributed a volume by the late licentiate of theology from Erfurt, Johannes Carnificis de Lutrea, who had passed away three years earlier. The book is entitled *Exercitium librorum de anima,*[5] and it is probably one of the earliest printed commentaries on Aristotle's treatise *De anima*. After this publication, several books on natural philosophy were published by Bartholomaeus Arnoldi de Usingen and his elder colleague Jodocus Trutfetter.[6]

These treatises include several different types of publications. Some of these may be characterized as *compendia*, concise textbooks, in which the material is arranged in a more or less systematic manner. Such are Usingen's *Parvulus philosophie naturalis* (1499) and his later expositions on natural philosophy in the early years of the sixteenth century. The last one of these, called *Compendium philosophie naturalis* (1517), was reprinted as late as 1543 for the use of university teaching at Erfurt. Jodocus Trutfetter's *Summa in totam physicen* (1514) is a slightly lengthier exposition on natural philosophy, but the material is put together in a rather similar way to Usingen's *compendia*.

The textbooks include necessary terminological definitions and divisions on each topic, sometimes giving philosophers' various opinions on central matters or even discussing such opinions in an argumentative manner. Trutfetter's *Summa in totam physicen*, which was first published in 1514, actually collects and sums up the whole tradition of medieval natural philosophy to the most recent authors, trying to formulate the position of the *via moderna* in every single question.

---

antiqua and *via moderna* in the Fifteenth Century: Doctrinal, Institutional, and Church Political Factors in the *Wegestreit*," in L. Nielsen & R. Friedman (eds.), *The Medieval Heritage in Early Modern Metaphysics and Modal Theory* (Dordrecht: Kluwer, 2003).

[3] Kleineidam, *Universitas II*, 140–45; Götz-Rüdiger Tewes, "Die Erfurter Nominalisten und ihre thomistischen Widersacher in Köln, Leipzig und Wittenberg. Ein Beitrag zum deutschen Humanismus am Vorabend der Reformation," in Andreas Speer (ed.), *Die Bibliotheca Amploniana* (Berlin: Walter de Gruyter, 1995).

[4] Kleineidam, *Universitas II*, 260–66.

[5] Johannes Carnificis de Lutrea, *Exercitium librorum de anima* (Erfurt: Paulus Wider de Hornbach, 1482).

[6] For their lives and works, see Kleineidam, *Universitas II,* 153–57; 312.

The relationship between these textbooks and the actual lectures on Aristotelian natural philosophy is not clear, but it is rather obvious that the lectures given at Erfurt were not completely dependent on the commented text; this gave the lecturer freedom to concentrate on what were regarded as the central topics. For this kind of lecturing, these textbooks would have been of great help.[7]

In addition to such textbooks, there was a distinct type of publication called the *exercitium*. It consisted almost entirely of independent questions, arranged according to the books and chapters of Aristotle's texts. Besides Lutrea's *Exercitium in libros de anima* from 1482, there is also *Exercitium de anima* by Usingen from 1507.[8] Their titles suggest that they were used in the disputation exercises, which formed a fixed part of the curriculum. Whereas the textbooks presented a rather large variation in contents, the *exercitia* followed a strict formal structure like that of the earlier question commentaries.

These *exercitia* belong to a line of tradition initiated by John Buridan's question commentaries.[9] In comparison to Buridan's commentaries, they are considerably economical in their wording, but at the same time they are abundant in the number of counter-arguments and their solutions. This applies especially to Usingen's *Exercitium de anima*, which seems to be based on Lutrea's commentary. The questions follow a regular pattern: the title of the question is followed by a number of *notabilia*, where necessary terminological divisions and definitions are made; the response to the question comes after *notabilia*, often in the form of one or two conclusions and their proofs; after that come the objections to the response together with their solutions.

Lutrea's and Usingen's *exercitia* contain the most elaborate discussions concerning questions of philosophical psychology at the late medieval University of Erfurt. Their arguments derive from an accumulated tradition of Buridanian natural philosophy, collecting as large a number of arguments as possible and often presenting such extensive terminological divisions that they far exceed the needs of the question at hand. With such a high standard of sophistication, it is no wonder that there were considerable difficulties in continuing the inherited tradition of disputation exercises at the University of Erfurt in the early sixteenth century. These problems were not made easier by the emerging humanist reforms in studies from the second decade

---

[7] Indications of such a use of textbooks can be found in Trutfetter's printed commentaries on the lectured books of logic in which the author continuously refers to his own textbooks. See, for example, Jodocus Trutfetter, *Veteris artis: id est Porpyrii universalium: et praedicamentorum Aristotelis: Perihermeneiasque expositio tam brevis quam utilis una cum dubiolorum ex iis extractorum resolutione* (s.l., s.a.).

[8] Bartholomaeus Arnoldi de Usingen, *Exercitium de anima* (Erfurt: Wolffgang Schenck, 1507).

[9] For the text of Buridan's commentaries see Peter Gordon Sobol, *John Buridan on the Soul and Sensation. An Edition of Book II of His Commentary on Aristotle's Book on the Soul with an Introduction and a Translation of Question 18 on Sensible Species* (Diss. Indiana University, 1984) and John Alexander Zupko, *John Buridan's Philosophy of Mind: An Edition and Translation of Book III of His 'Questions of Aristotle's De Anima' (Third Redaction), with Commentary and Critical and Interpretative Essay* (Diss. Cornell University, 1989); for the Buridanian tradition see Peter Marshall, "Parisian Psychology in the Mid-Fourteenth Century," in *Archives d'histoire doctrinale et littéraire du moyen âge* 50 (1984), 101–93.

of the sixteenth century onwards, neither by the general decline of the teaching at the university in the early 1520s. So at the peak of its development, the form of the question commentary gradually became useless.[10]

2.1.2 The Characteristics of Nominalist Philosophical Psychology at Erfurt

It is impossible to describe here the details of the *via moderna* method of reading Aristotle's *De anima* at Erfurt, but it is useful to point out some general doctrines that were regarded as especially important to the nominalist teachers as regards the way they understood their doctrinal position. The university more or less normatively prescribed some of these as positions to be affirmed in teaching.[11] The first is the rejection of the idea that there is more than one substantial form in compounds of form and matter. This meant an explicit rejection of Scotus' and Ockham's theories which posit a distinct form of corporeity, or even, as in the case of Ockham, a distinct sensitive soul alongside a rational soul in human beings. The teaching at Erfurt followed Aquinas's and Buridan's solution to this matter.[12]

Regarding the soul, it was thought that there were no real differences among its various potencies or faculties. Moreover, such terms as reason, will, and sensitive potency were deemed only as names that stood for the soul itself, connoting different types of acts it produced. This doctrine was, unlike the former, directed against the Thomist tradition, in which potencies of the soul were distinguished differently.[13] Similarly, the Erfurt teachers of the *via moderna* held the conviction (against the Thomist tradition of conceptual realism) that there were no universal entities (*universale in essendo*), that all entities were according to their nature singular and that universals existed only as concepts in the mind.[14]

---

[10] See Kleineidam, *Universitas II,* 161; 191; 205.

[11] Kleineidam, *Universitas Studii Erffordensis I* (Leipzig: St. Benno-Verlag, 1964), 181–91.

[12] See Kleineidam, *Universitas I,* 185; Lutrea, *Exercitium,* 21^r: "Respondetur ad questionem quod in uno homine etiam una forma substantialis et non plures qui claudit in se omnes perfectiones omnium inferiarum formarum. Illa responsio est beati Thome Marsilii Byridani Aristotelis in hoc secundo de anima"; Usingen, *Parvulus philosophie naturalis* (Leipzig: Wolfgang Stöckel, 1499), 88^r; *Ex. an.* D3^v; *Compendium,* K3^r; Jodocus Trutfetter, *Summa in totam physicen*, (Erfurt: Mathias Maler, 1514), Y1^v.

[13] Lutrea, *Ex.* 23^r–26^r; Usingen, *Parv.* 82^v–84^v; *Ex. an.* E1^v–E5^r; *Compendium naturalis philosophie* (Erfurt, Wolffgang Schenck, 1505–7) K4^r; Trutfetter, *Summa* Y1^r; X3^r–4^r. The Thomist position defined the potencies as qualities or properties of the soul. They emanate from the essence of the soul and are really distinct from its essence as well as from each other. For a contemporary example of a Thomist position, see Johannes Peyligk, *Philosophiae naturalis compendium* (Leipzig: Melchior Lotter, 1499), K1^v–K2^r.

[14] Lutrea, *Ex.* 11^v; Usingen, *Ex. an.* B4^v. See also *Questio preambularis in disputacionem de quolibet celebratam Erffordie A. D. 1497,* I.1.3, ed. Hans-Ulrich Wöhler in *Bochumer philosophisches Jahrbuch für Antike und Mittelalter* 6 (2001), 152: "Nulla est entitas universalis respectiva, indivisibilis, privativa aut successiva". On the importance of the quodlibet, see Wöhler's preface to the edition.

The human soul was understood to be both a substantial form of the body and an immaterial, incorruptible, and unextended entity.[15] Even before the famous decree of the Fifth Lateran Council on the immortality of the soul, it was taught that the immortality of the soul is true not only in theology but also in natural philosophy. The peculiar character of the *via moderna*, however, can be observed in Usingen's conception that this truth can be proven in natural philosophy only on the basis of probable reasons, not by an evident demonstration.[16] The decree of the Fifth Lateran Council on immortality was known to Jodocus Trutfetter when he wrote his *Summa in totam physicen*, published the year after the Council's decree. Trutfetter also seems to have applied the order of the council: the immortality of the soul should be vindicated in the teaching of natural philosophy.[17]

2.1.3 Some Individual Themes

Before the publication of Gabriel Biel's theological commentaries, around the year 1500, the teaching of philosophical psychology mostly relied on the tradition based on Buridan's commentaries of *De anima*. Among Buridan's followers, such authors as Peter of Ailly and Lawrence of Lindores also had some impact on teaching at Erfurt. These and some other authorities, like Gregory of Rimini, supplied a discussion and a set of solutions that different writers at Erfurt valued to different degrees, though the differences in their positions may seem rather trivial to a modern reader.[18]

Biel's overall influence was striking within the writings of sixteenth century philosophical psychology at Erfurt. This may be seen, for example, in the different evaluations regarding the theory of sensible species. According to the standard theory, shared by Buridan, species are defined as forms or likenesses of the objects of sensation. They are produced by the objects and multiplied in the intermediating substance to produce similar species in the sense organs. These likenesses are said to be not sensible in themselves but needed to produce the act of sensation in the soul. Another kind of explanation was given by William Ockham in his *Sentences* commentary, and, after Peter of Ailly, it was presented as a possible solution to the question of how objects cause the act of sense perception. According to this view, objects produce acts of sensation in the sense organs by means of an action at a distance, without any intermediating *species sensibiles*. Furthermore, objects produce sensible qualities that are the same in kind as their own, and these qualities

---

[15] Lutrea, *Ex.* 54^r; Usingen, *Parvulus,* 112^{r-v}; *Ex. an.* L3^r–4^r; Trutfetter, *Summa* Y2^v. The quodlibetal dispute in 1497 already addressed the question of the origin of the human soul among other topics, reaffirming the Catholic doctrine of the creation and immortality of individual souls. See *Questio preambularis,* I.3.1 (ed. Wöhler, 154); Trutfetter, *Summa,* P4^r.

[16] Lutrea, *Ex.* 58^v; Usingen, *Ex. an.* M4^v. See also below, n. 34.

[17] Trutfetter, *Summa* Y2^v–Y3^r; Y4^r.

[18] Lawrence of Lindores is mentioned both by Lutrea and Usingen. See Lutrea, *Ex.* 38^v; Usingen, *Ex. an*. H4^v. Both Usingen and Trutfetter refer frequently to Peter of Ailly, Gregory of Rimini and Biel.

arise both in the intermediating substance and also in the sense organs; however, these sensible qualities are not required for an act of sensation. The necessary conditions for an act of sensation consists, according to Ockham, merely of two things: first, the object must be present, and, second, there must be a sense organ that is properly disposed to sensation.[19]

The Erfurt authors seem to be open to both Buridan's and Ockham's views, conflicting as they are. Usingen notes in his *Parvulus philosophie naturalis* that Ockham's position has not found much support, whereas the contrary position that he describes as the common position of the *via moderna*, is held by Peter of Ailly, Buridan, Marsilius of Inghen, and "by almost all the philosophers, because only few or nobody follow Ockham in his opinion."[20] Trutfetter seems to be influenced more by the strictly Ockhamist Sentences commentary of Biel and is less sure of the validity of the traditional view; however, he does present the entire process of sense perception by using the notion of sensible species.[21]

The problems concerning the five individual senses centre on the different ways the sensible species and other involved qualities are transmitted by multiplication in a mediating substance, like air. The authors hold that light (*lux*) is perceived through the species of light and luminosity (*lumen*), a quality in the medium caused by light, by which the medium is called luminous. Both species of light and qualities of luminosity are multiplied instantaneously, without any succession of time if the medium is perfectly transparent.[22]

As a standard question among the Buridanian *via moderna* it is inquired here whether luminosity is needed only to make colours visible by creating corresponding sensible species or whether it is needed to illuminate the medium, which is the same as to make the medium transparent. According to Lutrea, luminosity is needed only to make colours visible. Stating this, he explicitly follows the interpretation of the Arabic philosopher Avempace (Ibn Bajja) shared by Lawrence of Lindores and found in Buridan's final version of his commentary on *De anima*.[23] In *Parvulus philosophie naturalis*, Usingen seems to favour Lutrea's position, calling it most probable. Later in his *Exercitium de anima*, Usingen disagrees with Lutrea on this point: according to him, luminosity is needed both to make colours visible and to illuminate the medium.[24]

---

[19] See Anneliese Maier, "Das Problem der 'species sensibiles in medio' und die neue Naturphilosophie des 14. Jahrhunderts," in *Ausgehendes Mittelalter II* (Rome 1967), 419–51 for the history of the discussion.

[20] Usingen, *Parvulus* 85[r].

[21] Trutfetter, *Summa,* Z6[v]; Aa1[v]; Aa2[r].

[22] Lutrea, *Ex.* 36[v]; Usingen, *Parvulus* 100[r]; *Ex. an.* H3[v].

[23] "Alia est opinio Avenpote et Londorii et Aristotelis quodammodo. Et est quod lumen requiratur propter colores et non propter medium. Et secundum illam opinionem respondetur ad questionem quod ad videndum colores requiratur lumen propter colores et non propter medium". Lutrea, *Ex.* 38[v]. See John Buridan, *Quaestiones in tres libros De anima Aristotelis* (Tertia lectura), II, 15, trans. in Sobol, *John Buridan,* 222.

[24] Usingen, *Ex. an.* H2[v]. Also, Trutfetter discusses the question and he agrees with Usingen's final position. See Trutfetter, *Summa* Bb2[r].

The sensible species are thought to be things that reach the observer by multiplying themselves in the medium. The species of light was thought to multiply instantaneously, whereas the species of sound to multiply successively in time.[25] On the nature of sound itself, Usingen presents Ockham's position, also found in Biel, that sound is a mere fraction of air. He does not judge between this and the traditional explanation, that sound is a distinct quality that comes about by a fraction of air; he notes that the question "remains hidden in the secrets of nature."[26] Trutfetter discusses the problem in his commentary on *De sensu et sensato*. He favours the traditional view, noting that it is more common.[27]

Concerning the transmission of odour in the medium, Lutrea argues for what was understood as Averroës' position. According to this, odour is multiplied intentionally as *species odoris* and not in a real way by qualities that may be sensed as odour.[28] Usingen follows Buridan's harmonization of Avicenna's and Averroës' positions: odour multiplies both as a real, perceptible quality of odour and as *species odoris* which is not perceived in itself.[29]

Discussions like the ones described above were not exceptional in the writings on philosophical psychology at Erfurt. It is surprising how much of the discussion continued around the same questions as those discussed since the introduction of the Arabic tradition of Aristotle's commentaries in Western philosophy. This applies not only to the above-mentioned questions concerning the sensitive soul but also to reason and its operations.

On the nature of the intellect, the questions raised by the confrontation of Christian theology with Averroës' commentary on *De anima* were still dominant. These included such problems as how the intellectual soul can be at the same time separate from the body and the substantial form of the body, questions about the immortality of the soul, and whether there is one common intellectual soul in all human beings.[30]

---

[25] "Et differunt species coloris et soni. Primo quia species soni multiplicatur instantanee sed species soni successive, quia species soni subiective sunt in aere". Lutrea, *Ex.* 40ʳ; see also Usingen, *Parv.* 100ʳ; *Ex. an.* I1ʳ.

[26] Usingen, *Ex. an.* H4ᵛ–I1ʳ. See also Usingen, *Parv.* 98ᵛ–99ʳ. Lutrea defines sound as quality that is nothing other than a mere fraction of the air, and which inheres in the air as its subject. "Notandum est quod sonus sic describitur est qualitas secunda ex corpore concutiente et concussioni aere incepto generata alterativa auditus de potentia ad actum". Lutrea, *Ex.* 40ʳ; "Respondetur quod sonus sit formaliter in aere tamquam in subiecto quia sonus nihil aliud est nisi fractio aeris sed talis fractio in nullo alio est nisi in corpore fracto". 40ᵛ; see also 47ʳ.

[27] Trutfetter, *Summa* Ii3ʳ.

[28] Lutrea, *Ex.* 41ᵛ. For the meaning of intentionality in this context see Catherine H. Tachau, "Some Aspects of Intentional Existence at Paris 1250–1320," in S. Ebbesen and R.L. Friedman (eds.), *Medieval Analyses in Language and Cognition*, Acts of the symposium The Copenhagen School of Medieval Philosophy January 10–13, 1996, Historisk-filosofiske Meddelelser 77 (Copenhagen: The Royal Danish Academy of Sciences and Letters, 1999), 331–53.

[29] Usingen, *Parv.* 101ᵛ; *Ex. an.* I2ʳ; *Comp.* M2ʳ; Trutfetter, *Summa* Ff1ʳ–Ff5ᵛ.

[30] Already, Albert the Great had attacked Averroës on these matters, and, later, Thomas of Aquinas and Bonaventure identify the contemporary Averroists by their positions on these controversial issues. See Katherine Park, "Albert's Influence on Late Medieval Psychology," in James A. Weisheipl

The position of the Erfurtians on these matters was fixed by the longstanding Buridanian tradition of the university. They taught that the human intellect was incorruptible and does not perish with the body, as the famous ancient peripatetic philosopher Alexander of Aphrodisias had maintained. Furthermore, they rejected Averroës' doctrine, according to which all human intellects are numerically one; they taught instead that all human beings have their own individual intellectual souls. These intellectual souls were considered as substantial forms of the individual bodies inhering in them and not merely assisting them as Averroës had thought.[31]

The Erfurtians taught that the immortal human intellectual soul was responsible for all the functions that were attributed to the passive and agent intellect in Aristotelian psychology.[32] Therefore, there was no need to posit God as the agent intellect, as Alexander of Aphrodisias and many others, including Buridan, had done. As a modification of Alexander's position, Buridan had argued that both God and the human soul co-operate as one agent intellect.[33]

Despite this and other differences, the Erfurtians followed for the most part the positions on the human intellectual soul held by Buridan in his commentaries on *De anima.* One difference between Buridan and his late medieval followers in Erfurt was that the latter emphatically defended the philosophical nature of some doctrines concerning the intellectual soul; however, for Buridan the plausibility of these doctrines in natural philosophy seemed suspect. Even if the Erfurt Buridanians tried to avoid the accusation of holding the doctrine of double truths, viz., the idea that a statement can be true in theology and false in philosophy, not all of them, however, adopted very fervent ideas about the philosophical provability of the theological doctrine that human souls are at the same time both immortal and substantial forms of bodies. Only Lutrea seems to have adopted the view that this doctrine can be proven in natural philosophy without any specifications. Usingen, however, declared that the philosophical arguments for this doctrine do not yield a naturally demonstrated truth but only a dialectically probable truth.[34] By holding this middle view between the more radical Averroists and Alexandrists on one hand and their Thomist adversaries on the other, they were able to avoid the controversies that plagued many other universities of the time, most notably the universities of Padua and Bologna in Italy.

(ed.), *Albertus Magnus and the Sciences. Commemorative Essays 1980*, Studies and Texts vol. 49 (Toronto: Pontifical Institute of Medieval Studies, 1980), 511.

[31] See above, n. 13.

[32] Lutrea, *Ex.* 60ʳ; Usingen, *Ex. an.* N1ʳ;2ᵛ; see also Usingen, *Parv.* 116ᵛ; *Comp.* M5ʳ⁻ᵛ.

[33] Lutrea, *Ex.* 61ᵛ; Usingen, *Ex. an.* N2ᵛ. On Alexander and Buridan, see Maarten J. F. M. Hoenen, *"Die Intellektlehre des Johannes Buridan – ihre Quellen und historisch-doktrinären Zügen,"* in E.P. Bos & H.A. Krop (eds.), *John Buridan: A Master of Arts. Some Aspects of his* Philosophy, Artistarium Supplementa VIII (Nijmegen: Ingenium 1993), 89–106.

[34] Pekka Kärkkäinen, "Nominalist Psychology and the Limits of Canon Law in Late Medieval Erfurt," in Virpi Mäkinen (ed.), *Lutheran Reformation and the Law* (Leiden: Brill, 2006), 93–110.

2.2 Philosophical Psychology at the Universities of Padua and Bologna

The Italian universities in Bologna and Padua are not only the oldest in Italy, but also, around 1500, they were two of the most important institutions in Europe, where Italy's most important philosophers and thinkers were located. Padua, however, stands out as the foremost university in Italy of the time with a long-standing tradition in philosophy. In 1509, the Venetian army suffered a devastating loss and in the turbulence that followed the university lost most of its students. As a result many teachers also left. Among them was Pietro Pomponazzi (1462–1525) – perhaps the most famous philosopher of the time. He then moved to Bologna. During the period under consideration here, however, Padua was the central institution for philosophy in Italy.

The "School of Padua" and "Paduan Aristotelianism" are familiar labels in Renaissance philosophy. This strong and influential tradition, foremost in natural philosophy, stretches over the three centuries from Pietro d'Abano (1250–1316) to Jacopo Zabarella (1533–1589). During this time, there was a very close relationship between the universities in Padua and Bologna, thus the school should really be known as the "School of Bologna and Padua".

The school was, however, never very homogeneous, but was instead made up of some loosely connected philosophers, active as teachers at both Padua and Bologna. These philosophers were all writing and teaching in a scholastic tradition, and, hence, they all belonged to a broadly Aristotelian tradition. We write "broadly" because their Aristotelianism was coupled with a variety of different sources and because it has been labelled "Eclectic Aristotelianism".[35]

As will become evident, and as has been emphasized by numerous books and articles written about the topic, this school had a particular interest in, and focus on, natural philosophy. The relative outspokenness of some of the philosophers in this school can to some extent be explained by the fact that theology had a rather weak position in Bologna and Padua. Instead, the connection between natural philosophy and medicine was emphasized. The theological authorities, of course, sometimes tried to intervene when Aristotelianism, foremost in psychology, became too extreme, but these interventions were never very serious and they did not stop anybody from teaching what they wanted.

Besides Aristotle, the main authority for many philosophers in this school was Averroës. He was consistently influential during most of the fifteenth century, but around 1500 his influence had somewhat declined. Averroës was still important, but the translation of several Greek commentators on Aristotle had changed the school to some extent.[36]

---

[35] See C.B. Schmitt, *Aristotle and the Renaissance* (Cambridge, Mass.: Harvard University Press, 1983), 99–103.

[36] Historians of science have studied the Padua School to a particularly large extent. See P.O. Kristeller, "Paduan Averroism and Alexandrism in the Light of Recient Studies," in *Aristotelismo padovano e filosofia aristotelica, Atti del XII Congresso Internazionale di Filosofia, IX,* (Florence,

2.2.1 The Teaching of Natural Philosophy at Padua and Bologna

The teaching of philosophy was divided in the usual manner at both Padua and Bologna, but the emphasis was on logic and natural philosophy. At both universities, teaching in natural philosophy was considered fundamental, and it was perhaps emphasized more strongly than at other European universities, since natural philosophy was considered to be preparatory to medical studies. This strong relation between natural philosophy and medicine was a heritage that came particularly from Paduan authors like Pietro d'Abano in the early fourteenth century.

The core of the curriculum was, of course, the natural philosophy texts of Aristotle. The statutes of the University of Padua have unfortunately been lost, but those of the University of Bologna from 1405 show that studies usually began with *Physics* and book 1 of *De generatione et corruptione*. The second year was devoted to *De caelo et mundo*, *Meterology* and *De sensu et sensibili*. It was only during the third year that students heard lectures on *De anima*[37] and *De motu animalium*. During that year the course on natural philosophy also included lectures on the *Metaphysics*. The emphasis on *De anima* seems to have grown over the years, and soon the first year students heard lectures on all three books of *De anima* as well.[38]

Towards the late fifteenth century, some of the late ancient Greek commentators of Aristotle had been translated into Latin, and they gradually began to find their way into the curriculum at these two universities. One of the first of these commentators was Simplicius, and, later, Philoponus' commentaries were also used for teaching. A novelty was that some began to use the Greek text of Aristotle as basis for their teaching – although it was not a very widespread technique.[39]

In general, a wide variety of sources on natural philosophy were used to facilitate the professors in their teaching, including Aristotle, Aristotle's Greek commentators,

---

1960), 147–55; Bruno Nardi, *Saggi sull'aristotelismo padovano del secolo XIV al XVI* (Florence, 1961); J.H. Randall, *The School of Padua and the Emergence of Modern Science* (Padua: Editrice Antenore, 1961); P.O. Kristeller, "The Myth of Renaissance Atheism and the French Tradition of Free Thought," *Journal of the History of Philosophy* 6 (1968), 233–43, Antonio Poppi, *Introduzione all'aristotelismo padovano* (Padua, 1970), E.P. Mahoney, "Saint Thomas and the School of Padua at the End of the Fifteenth Century," in *Proceeding of the American Catholic Philosophical Association: Thomas and Bonaventure. A Septicentenary Commemoration*, vol. 48 (1974), 277–85; J.H. Randall, *The Making of the Modern Mind* (New York: Colombia University Press, 1976); *ibid.* "Paduan Aristotelianism Reconsidered," in E.P. Mahoney (ed.), *Philosophy and Humanism: Renaissance Essays in Honor of Paul Oskar Kristeller* (New York and Leiden, 1976), 275–82; E.P. Mahoney, "Albert the Great and the Studio Patavino in the Late Fifteenth and Early Sixteenth Centuries," in J.A. Weisheipl (ed.), *Albertus Magnus and the Sciences* (Toronto: Pontifical Institute of Mediaeval Studies, 1980), 537–66; Antonio Poppi (ed.), *Scienza e filosofia all'Università di Padova nel Quattrocento* (Trieste: Lint, 1983) and Schmitt, *Aristotle*.

[37] In Carlo Malagola (ed.), *Statuti delle universitá e dei collegi dello Studio bolognese* (Bologna, 1966), 274, it is said that the professors of natural philosophy were not allowed to lecture on the whole of *De anima*, since certain "errors" were to be left out. It is not said what errors they were.

[38] See Malagola, *Statuti* and Paul F. Grendler, *The Universities of the Italian Renaissance* (Baltimore and London: The Johns Hopkins University Press, 2002), 269–71.

[39] See Grendler, *The Universities*, 271–77.

Plato, the Arabs (foremost Averroës of course), and some later medieval commentators. It is this mix of sources that lies behind the term "Eclectic Aristotelianism", but it is of course important to stress that Aristotle was still the main authority.

2.2.2 The Philosophers and Their Writings on Philosophical Psychology

The trends in philosophical psychology at Padua and Bologna were set by two Italian philosophers, namely, Blasius of Parma (1347–1416) and Paul of Venice (1369–1429), though they were educated elsewhere. Blasius of Parma obtained his doctorate from the University of Paris and was heavily influenced by the thinking of John Buridan and Marsilius of Inghen. Blasius moved back to Italy at the beginning of the fifteenth century and became a professor of philosophy at Padua in 1403. His foremost work on philosophical psychology is called *Quaestiones de anima* and exists in several redactions. The first redaction, from 1385, was written in Padua.[40]

Paul of Venice, who was educated at Padua and Oxford, held teaching positions at both Padua and Bologna. He is foremost known as a logician and his textbooks on logic were widely used throughout the fifteenth century. In 1408, he produced a great compendium called *Summa philosophiae naturalis*, containing a discussion of *De anima*; moreover, he later composed a commentary on Aristotle's *De anima*, sometime between 1415 and 1420. Paul of Venice was the principal instigator of the great Averroistic trends in psychology at Padua and Bolognia in the fifteenth century.[41]

Another important philosopher at Padua in the early fifteenth century was Gaetano da Thiene (1387–1465). He became professor of natural philosophy at Padua in 1430 and produced several works and commentaries. Gaetano was succeeded by one of his own students, Nicoletto Vernia (1420–1499). Vernia was a prominent figure in Paduan Aristotelianism and a devoted Averroist until late in life when he was forced to give it up. The Bishop of Padua, Pietro Barozzo, issued in 1489 the *Decretum contra disputants de unitate intellectus*, in which he forbade all further public discussion of the Averroist doctrine of the unity of the intellect.[42] Vernia was the target of this attack. In his later book against the Averroist doctrine, *Contra perversam Averrois opinionem*, Vernia argues against Averroës and the

---

[40] See Graziella Frederici Vescovini, *Le Quaestiones de anima di Biagio Pelacani da Parma* (Florence: Leo S. Olschki Editore, 1974) and "Blasius of Parma," in E. Craig (ed.), *Routledge Encyclopedia of Philosophy* (London: Routledge, 1998) for a more detailed account of his life and works.

[41] See Nardi, *Saggi sull'aristotelismo padovano*.

[42] See Francesco Scipione Dindi Dall'Orologio, *Dissertazioni sopra l'istoria ecclesiastica di Padova: Dissertazione nona* (Padua 1817), 130–31, and Pietro Ragnisco, *Documenti inediti e rari intorno alla vita ed agli scritti di Nicoletto Vernia e di Elia del Medigo* (Padua 1891), 8–9. The decree was not intended to cover classroom lectures.

fourteenth century Averroist John of Janduno.[43] He cites a great number of authors in this book, including Alexander of Aphrodisias and several other late ancient commentators of Aristotle.[44]

Antonio Trombetta (1436–1517) was a representative of what is often called the *via scoti* at Padua around 1500. In the work *Tractatus singularis contra Averroistas de humanarum animarum plurificatione ad catholicae fidei obsequium Patavii editus* (Venice, 1498), he heavily attacks the Averroists at Padua. He thinks that they have wrongly interpreted Averroës and consequently tries to explain what Averroës' position really is. He thereby uses much of Averroës' thinking to develop his own position. An example is his defence of a pluralist view of the human soul.

Alessandro Achillini (1463–1512), influenced by William Ockham, was responsible for a revival in the interest of Ockham's thinking at Bologna around 1500; otherwise, Ockham seems not to have been influential at all at Padua.[45] Achillini was a professor of philosophy at both Bologna and Padua. His most important work in psychology is *Quaestiones quodlibeta de intelligentiis* from 1494.[46]

A well-known teacher at Padua around 1500 was Jacopo Vio (1468–1534), better known as Thomas Cajetanus. In order to reinforce the position of Aquinas against the Averroists at Padua, a special post in metaphysics was created. Cajetanus held that chair between 1497 and 1499. His main writings on psychology are his commentary on Aquinas' *Summa theologiae* and his *De anima* commentary from 1510.

Agostino Nifo (1470–1538), though less well known than Cajetanus, was a prominent figure at the Padua School. He was a student of Vernia and, therefore, got to know the texts of Averroës early. He translated and wrote a commentary on Averroës' *Tahafut al-Tahafut*, called *Destructio destructionum*, which was published in Venice in 1497. His most important work in psychology is, however, *De intellectu*.

The last member of the Padua School that we will mention is Pietro Pomponazzi (1462–1525). He is, of course, the most famous of the philosophers mentioned here, and this is precisely because of his views on the immortality of the human soul. He was also a student of Vernia. His most renowned work is *De immortalitate animae* from 1518, but during his 18 years as a professor at Padua he wrote a large number of works, many of which were never printed.[47]

---

[43] See Nicoletto Vernia, *Contra perversam Averrois opinionem de unitae intellectus et de animae felicitate questiones divinae*, Venice 1505, fol. 7a.

[44] See also E.P. Mahoney, "Nicoletto Vernia and Agustino Nifo on Alexander of Aphrodisias: An Unnoted Dispute," *Rivista Critica di Storia della Filosofia* (1968), Fasc. III, 270–71.

[45] See John Monfasani, "Aristotelians, Platonists, and the Missing Ockhamists: Philosophical Liberty in Pre-Reformation Italy," *Renaissance Quarterly* 46 (1993), 247–76 about the absence of Ockhamism in Italy.

[46] See Herbert Matsen "Alessandro Achillini (1463–1512) and 'Ockhamism' at Bologna (1490–1500)," *Journal of the History of Philosophy* 13 (1975), 437–51 for a study of Achillini's Ockhamism. Achillini seems to have been closer to Averroës' position when it comes to human nature and cognition; his Ockhamism is limited to universals and other positions in metaphysics.

[47] See Poppi, *Introduzione*.

The so-called Bologna and Padua School, of course, consists of many more philosophers than have been mentioned here, but we think that this is a representative selection. It is also apparent from what has been said so far that the "school" is hardly homogeneous but consists of a number of different views on human nature and cognition. The most salient characteristic is, of course, the Averroism that was prominent at least from Paul of Venice onward but can be traced back as far as Pietro d'Abano. During the time period under consideration, we thus find representatives of Averroism, Alexandrianism, Scotism, Thomism and, to some extent, Ockhamism. To flesh out these labels somewhat, let us have a look at some of the main issues that were debated within this school.

2.2.3 The Main Issues Discussed

Philosophical psychology within the Padua School was, as mentioned, thoroughly Aristotelian, and as such it should be seen as a continuation of a long scholastic tradition. The issues that were treated and debated were therefore virtually the same as those that had existed within the tradition for over three hundred years, since the translation of *De anima* into Latin. Some issues, though not new, had become increasingly controversial. One such issue had to do with the nature of the human soul. The central concern was how to reconcile Aristotle's view of the relation between body and soul with traditional dogmas within the Christian faith, particularly the soul's immortality.

The height of this controversy was reached with the decree of the Fifth Lateran Council and the so-called Pomponazzi affair. Pomponazzi famously argued that the immortality of the human soul could not be proven by natural reason. He therefore ended his famous treaties with a doubt. He wrote: "For it seems to me that no natural reasons can be brought forth proving that the soul is immortal, and still less any proving that the soul is mortal, as very many scholars who hold it immortal declare. Wherefore I do not want to make answer to the other side [. . .]."[48] Despite this acknowledgment, Pomponazzi was not only criticized by the authorities but also by some of his colleagues within the faculty of philosophy. The most outspoken critique came from Nifo. He wrote the book *De immortalitate animae libellum* in 1518 as an answer to Pomponazzi's *De immortalitate animae*.

Pomponazzi defended himself with the claim that he was only doing natural philosophy and did not want to get involved in the theological debate. The Fifth Lateran Council had, however, already decided that this was not an acceptable excuse and issued a decree in 1513 that the immortality of the human soul should be taught in natural philosophy as well.

---

[48] Pietro Pomponazzi, "On the Immortality of the Soul," translated by W.H. Hay, in E. Cassirer, P.O. Kristeller and J.H. Randall Jr. (eds.), *The Renaissance Philosophy of Man* (Chicago: The University of Chicago Press, 1998), 377.

If one takes a closer look at what Pomponazzi actually writes, it soon becomes apparent that the authorities were right to be concerned. In Chapter 2 of his treaties, he divides his study into six groups. First, he notes that since mortal and immortal are opposite attributes they cannot be attributed to the same nature. Hence, they must be attributed to different natures and this can be done in three ways, namely:[49]

   i. The number of mortal and immortal natures will be the same as the number of humans.
  ii. Each human will have his or her own mortal nature and take part in a common immortal nature.
 iii. Each human will have his or her own immortal nature and take part in a common mortal nature.

Reflecting on these three alternatives we see that (i) corresponds to a position that assumes a plurality of souls (or natures) in each human being. This was a position defended by several Franciscan thinkers in the late thirteenth and early fourteenth century, among whom were John Duns Scotus and William Ockham (see also the discussion of the situation at Erfurt in the first part of this chapter). The second view corresponds to Averroës' position, namely, that the body and the sensitive soul are in this life united to a universal and divine intellect. Pomponazzi discusses (ii) very carefully, but dismisses it, and he does the same with (i) mainly for the same reason, namely, that a distinction between a sensitive and intellective soul cannot be maintained.[50] He does not discuss the third position and we know of no one who has held such a position.

If we, however, would like to hold on to the position that there is only one soul in each human, or one nature as he puts it, then we have to say that the soul is mortal or immortal, either *simpliciter* or *secundum quid*. Here he also comes up with three possible positions, namely:[51]

  iv. The soul is *simpliciter* mortal and *secundum quid* immortal.
   v. The soul is *simpliciter* immortal and *secundum quid* mortal.
  vi. The soul is *secundum quid* mortal and immortal.

He does not bother to discuss (vi). Nor does Pomponazzi deliberate on the distinction between *simpliciter* and *secundum quid*, but we will assume that he uses the distinction in the traditional manner.[52] Applied to the soul, the distinction implies

---

[49] See *ibid.*, 283–84.

[50] See *ibid.*, 284–99. See also Henrik Lagerlund, "John Buridan and the Problems of Dualism in the Early Fourteenth Century," *Journal of the History of Philosophy* 42 (2004), 369–87 for an outline of these positions in the fourteenth century.

[51] See Pomponazzi, "On the Immortality," 284.

[52] The distinction between *simpliciter* and *secundum quid* goes back to late medieval discussions of fallacies and is ultimately motivated by a distinction drawn by Aristotle himself in *Sophistici elenchi* 25 (180a26–31). Something can belong to something else without qualification or in a particular respect, and, hence, two contrary attributes can belong to the same thing if one belongs

that *simpliciter* immortal means that the soul is immortal absolutely or to its nature, but to say that it is also *secundum quid* mortal means that it also has functions that are typically mortal, for example, sensations and feelings.

If (v) is taken to mean that there are as many souls as there are humans, then (v) corresponds Thomas Aquinas' position; that is, the soul of each human is an immortal soul with mortal functions. Pomponazzi puts forward many doubts about this position, but he does not say "these things [. . .] to contradict so great a philosopher [. . .] but from desire of learning".[53] He is thus very cautious, particularly since this stance was considered to be that of the Catholic Church, but there is little room for doubt that (iv) is what he thinks natural reason demands one to accept, particularly since this is the only position for which he presents an explicit defence.

The view Pomponazzi thus seems to regard as the most plausible holds that the soul is mortal, that is, it dies with the body, but that it has functions that are immortal. This means that intellectual thought, in its functioning, abstracts from matter, since thinking is paradigmatically about universals while material things are particular individuals. This was not a new position – although it was, of course, controversial – and was often attributed to Alexander of Aphrodisias in the Middle Ages. It is one of the three views on the human soul outlined by Buridan in his commentary on Aristotle's *De anima* – a work that would have been know to most people in Padua through Blasius.[54] Blasius had himself argued that the soul is a naturally generated form and mortal.[55]

Pomponazzi's discussions of the soul thus have a very interesting and long medieval background, but his final position was by no means the only or even the most common opinion defended at the Padua School. In fact, all of the views that he discusses and refutes ((i), (ii) and (v), that is, the Franciscans', Averroës' and Aquinas' positions) had defenders within the school.

The discussion and defence of Averroës' views at the Padua School goes back, as already mentioned, to Paul of Venice. Around 1500, the most orthodox Averroist in psychology was probably Achillini. In his *Quodlibeta de intelligentiis* from 1494, he defends a position similar to the one Pomponazzi ascribes to Averroës. Achillini's source is, however, not only Averroës.[56] Another Averroist is Pomponazzi's main opponent, Nifo. He is, however, not an orthodox Averroist when it comes to the immortality of the soul, since he defends the view that there are as many immortal souls as there are bodies. He therefore takes the same position as his teacher, Vernia. He originally defends the same view as Paul of Venice and Achillini, but

---

to it without qualification and the other in a particular respect. In late medieval logic, the same distinction was used to point to the fallacy of confusing what is true in a certain respect with what is true absolutely. See, for example, P.V. Spade (1982) "Insolubilia," in N. Kretzmann, A. Kenny and J. Pinborg (eds.), *The Cambridge History of Later Medieval Philosophy*, (Cambridge: Cambridge University Press, 1982).

[53] See Pomponazzi, "On the Immortality," 313.

[54] See Lagerlund, *John Buridan* for a closer study of Buridan on human nature.

[55] See Federici Vescovini, *Le Quaestiones de anima*.

[56] See Nardi, *Saggi sull'aristotelismo padovano*.

after Barozzo's ban of Averroistic philosophy he changes his mind. Both Nifo and Vernia argue, contrary to Buridan and Pomponazzi, that Alexander does not actually defend the position that the soul is mortal.[57]

On the view Vernia defends later in life in the book *Contra perversam Averrois opinionem de unitate intellectus et de animae felicitate quaestiones divinae*, the soul is the substantial form of the body, it is created by God, and it is infused in the body. The soul is also multiplied and individuated in the body not through its relation to the body, however, as Aquinas took it to be, but in accordance with what Avicenna and Scotus call a "property" or "*haecceitas*" and what Vernia calls an "affection". This is the individuating condition of the person. He adds that there is only one soul in each living thing. He thinks these ideas can be shown by natural reason and that they can be shown to be the true views of Aristotle.[58]

Nifo and Vernia thus defend a stance that John Buridan had called the position of the Christian faith.[59] It implies a strong dualism between body and soul, and it is reminiscent of the position Descartes ends up with in the *Meditations*. This stance is not the same as Aquinas', namely, (v) in Pomponazzi's list, since it grants the soul a much more independent and self-subsistent status than Aquinas does. Aquinas, of course, had strong defenders at Padua at this time, the most influential being Cajetanus.

In book III of his commentary on Aristotle's *De anima*, Cajetanus includes, between paragraphs §103 and §141, a detailed discussion of the immortality of the soul. Here he claims that a proof of the soul's immortality can proceed in two ways, namely, through the soul's proper operations or through its potencies. With reference to *De anima* I, 1 (403a), he argues that if it can be established that there are operations of the soul that are completely independent of the body, then the soul is separable from the body. It is a very detailed and interesting discussion by which he thinks he has proven the immortality of the soul.

Another important issue debated at this school was the status of intelligible species. The school contained a variety of positions on this issue, as has been shown by Leen Spruit.[60] We can here only point to some interesting aspects of the discussions. One such aspect is a distinction traditionally associated with followers of Scotus, but which around 1500 had worked its way into mainstream philosophical psychology and had continued to be important all the way up to Descartes. This is the distinction between subjective, or formal, and objective existence.

---

[57] See Mahoney, *Nicoletto Vernia and Agustino Nifo* and Mahoney, "Antonio Trombetta and Agostino Nifo on Averroes and Intelligible Species: A Philosophical Dispute at the University of Padua," in A. Poppi (ed.), *Storia e Cultura al Santo di Padova fra il XIII e il XX secolo* (Venezia, 1976), 289–301.

[58] See E.P. Mahoney, "Nicoletto Vernia on the Soul and Immortality," in E.P. Mahoney (ed.), *Philosophy and Humanism: Renaissance Essays in Honor of Paul Oskar Kristeller* (New York and Leiden, 1976), 157–58.

[59] See Lagerlund, *John Buridan*, 383–84.

[60] See Leen Spruit, *Species Intelligibilis: From Perception to Knowledge,* Volume 2 (Leiden: Brill, 1995).

Cajetanus, who held a traditional Thomistic view on intelligible species, used the distinction to solve some problems in Thomas' theory of intellection. An important aspect of the Aristotelian or Thomistic theory is that intellection is like sensation.[61] In the same way as the senses take on the forms of the things sensed, the passive intellect takes on the form of the object of thought. It does this via the active intellect that turns towards the phantasm and abstracts the intelligible species. Cajetanus uses the mentioned distinction to elucidate how the actualization of the potential intelligible in the phantasm works.

The agent intellect illuminates the phantasm, and it abstracts the species objectively, not formally, meaning that it abstracts the general content and not the real particular thing as represented by the phantasm. He illustrates this through an analogy to sensation. Light shines through and illuminates the diaphanum of the eye formally, but the eye takes on the colour objectively, which is his explanation of why the eye does not become red when seeing something red.[62]

An interesting discussion of the intelligible species can also be found in Achillini's *De intelligentiis*. He argues that the species are not really distinct from the acts or the habits of the intellect and, hence, that there is only a distinction between them made by reason. The species is identified with the object of thought, but the object has a diminished kind of being in the intellect.[63] The advocated theory is quite reminiscent of Scotus' theory of cognition in which the intelligible species is retained but is seen as the content of the intellect's act. For Scotus as well as for Achillini, the act has subjective and formal existence, while the content has an objective and diminished kind of being.[64]

Nifo expresses a similar kind of view in his commentary on *De anima*. He very rarely uses the word *species*. Instead, he uses *intentio* and *notio*. Furthermore, these intentions or notions do not represent universals, but singulars. They are not subjective in the mind (*in mente*) but objective or intentional as mental objects.[65] Nifo

---

[61] See Thomas Cajetanus, *Commentaria in libros Aristotelis De anima III*, eds. G. Picard and G. Pelland (Paris: Desclée de Brouwer, 1965), III, §. 160.

[62] Cajetanus's doctrine of illumination is most clearly expressed in his commentary of Aquinas *Summa Theologiae*. See Spruit, *Species Intelligibilis*, 111–119, for a discussion of his theory.

[63] "De specie intelligibili actu intelligendi, aut habitu intellectus; doco quod sunt existentiae diminutae rerum quae sunt in de anima, sicut locatum in loco. [. . .] Quamvis igitur intellectus sit informatus accidentibus, non propter hoc illa sibi inhaerent, et haec informatio sine inhaerentia sufficit ad verificandum praedicationes accidentales et denominationes illorum de anima". (Alessandro Achillini, *Quaestiones quodlibeta de intelligentiis*, Bologna, 1494, 13va.)

[64] For Scotus's theory, see Robert Pasnau, "Cognition," in T. Williams (ed.), *The Cambridge Companion to Duns Scotus* (Cambridge: Cambridge University Press, 2003), 285–311.

[65] "At secundum veritatem per singula Aristoteles intelligit universalia, fit autem intellectus singula, quo recipit notiones, quae grece dicuntur noemata quae possunt dici species rerum intelligibilium. Ergo intellectus fit singula non realiter, ser intentionaliter, quatenus recipit singulorum species, mediantibus quibus quodammodo omnia est intentionaliter". (Agostino Nifo, *Expositio subtilissima collectanea commentariaque in III libros Aristotelis De anima*, Venice, 1553, 166vb.)

here shows himself as a follower of early Ockham or, more importantly, a follower of Peter of Ailly.[66]

2.3 Conclusions and Summary

Perhaps the most important conclusion we can draw from this preliminary study of philosophical psychology around 1500 is that it is surprisingly traditional. Firstly, medieval discussions on philosophical psychology carried on and developed the subject to its final maturity in the fourteenth century; this was done both in commentaries on Aristotle's treatise *De anima* and in treatises on problematic issues. Secondly, the results of studies were presented in textbooks that described different positions in a surprisingly clear and concise manner. This made it possible for students to gain relatively objective and systematic views of the discussions, instead of being taught only one approved position of their own school. Thirdly, the problems and questions related to the soul were virtually the same as those discussed in thirteenth and fourteenth century scholastic philosophy. Some problems had however become more significant than others. The problem of the nature of the soul and its purported immortality had, for example, a special status at this time and was consequently touched upon by all thinkers, even before the decree of the Fifth Lateran Council.

Another salient feature, which it seems important to stress, is that there was a surprising cross-fertilization between different traditions. Arriving at a coherent and defensible theory of the soul or cognition seems much more important for these thinkers than being doctrinally pure. For example, even though one followed Aquinas or Averroës in the main features of the theory defended, it was still important to use distinctions and arguments developed by thinkers in a rival tradition, such as in a Scotist tradition. This shows that the divisions between schools of thought were not very strict.

Let us end by noting a few aspects about the continuity between late medieval and early modern philosophical psychology. First of all, it is clear that much more needs to be understood about the details of philosophical psychology around 1500 to be able to make a full evaluation of our thesis. It seems, however, obvious that the problem of the nature of the human soul, which plays such an important role for Descartes and other seventeenth century thinkers, is a continuous problem from the later Middle Ages through 1500 and into later philosophy. Another over-arching aspect has to do with terminology. A distinction such as the one between formal and objective being was used in 1500 by all thinkers regardless of which school they belonged to. It is thus established as a useful distinction quite independent of its Scotist origin. We are quite certain that further study of this period will yield further aspects that will highlight the continuity we claim exists between these historical periods.

---

[66] See Henrik Lagerlund, "Representations, Concepts and Words – Peter of Ailly on Semantics and Psychology," *Proceedings of the Society for Medieval Logic and Metaphysics*, Vol. 3 (2003) for a discussion of Ailly.

Chapter 3
The Status of Psychology as Understood by Sixteenth-Century Scholastics

Tuomo Aho

Abstract The problems of the true nature and status of psychology (*scientia de anima*) among sciences were much debated by sixteenth-century philosophers. Chapter I.1 of Aristotle's *De anima* provided main material here, and different positions were mostly expressed in commentaries to it. This chapter discusses a number of sixteenth-century scholastic interpretations, in particular two great Jesuit works: the scholarly Coimbra commentary and Suárez's more original *De anima*. Aristotle himself posed several questions which required elaboration. One of these concerned the subject of psychology. Because of Averroism and Alexandrism, this issue was alive during the whole century. Another important problem was whether psychology belongs among natural sciences or to metaphysics. This generated a complicated discussion. Moreover, there was enquiry about what was the correct place of psychology among biosciences, how useful it was and how we should evaluate the worth and glory of it. A final puzzle concerned the explanation of its difficulty. The tentative conclusion of our survey is that these philosophers of the "second scholasticism" had a rather conscious notion of their task in this connection, though their methods were often tangled. Some of their observations, especially those related to the understanding of mental existence, can even have permanent relevance outside the strictly Aristotelian framework.

3.1 The Doctrinal Situation

In the late Middle Ages and the Renaissance, philosophers paid great attention to the system of sciences. Constantly interested in the harmony of the whole, they found pleasure in investigating the natures and correct characterisations of various disciplines. What we now wish to discuss, corresponding to the theme of our book, is the

T. Aho
Department of Philosophy, University of Helsinki

precise nature and role of the philosophy of the mind as a philosophical discipline, its relations to other parts of philosophy, and its choice of central problems – though we must almost bypass this last matter.

During the fifteenth and sixteenth centuries, the volume of philosophical literature grew explosively. In the philosophy of mind, much of this literature emerged in commentaries on Aristotle's *De anima*. The commentaries became more and more sophisticated. This was also connected to the publication of several revised text editions, new translations, and Greek commentaries.

From this period, we shall consider a set of books – well-known in their day – and the choice will concentrate mainly on the "second scholastics" of the latter half of the sixteenth century. Drawing clear historical boundaries is difficult here. The classical commentary tradition perhaps found a kind of summit in Thomas Cajetanus. Then, during the course of the sixteenth century, scholastic philosophy underwent considerable changes, but what really happened? That is controversial. Some scholars wish to underscore the continuity of the sixteenth century with older scholasticism, whereas others wish to emphasise the shift to new problems.[1] Indeed, at least in the philosophy of the soul, it is easy to see many new items.

Two important, that is substantially creative, currents were born in the latter half of the sixteenth century: the "second scholasticism" in Spain and Portugal, and the "scientific Aristotelianism" of Zabarella and others in Italy. In this chapter we will consider the philosophers of the former current with only a few comparisons to other Aristotelians.

In the late sixteenth century and in the beginning of the seventeenth century, Jesuits surely led scholastic philosophy. The Society of Jesus was founded without philosophers, but the second generation of its workers already included an important philosopher, Franciscus Toletus (1532–1596), and numerous others followed. The Jesuits took the philosophical initiative because of their general activeness and because they were not so bound to earlier schools. And they were also very interested in Aristotle's own works. For a long period, Jesuits authored the most important commentaries on Aristotle. The strictly Thomist Dominicans were much less active in this field.[2]

Thus my main subject will be two important Jesuit books, both because they are good books in themselves and because they are qualified representatives of "late Aristotelianism". They closely discuss the special nature and status of the science of soul. (In that science, such questions were exceptionally urgent because their solutions were anything but obvious.) These books are *Commentaria una cum quaestionibus in libros Aristotelis De anima*, by Francisco Suárez, and

---

[1] See Jorge Gracia, "Suárez (and Later Scholasticism)," in J. Marenbon (ed.), *Routledge History of Philosophy III: Medieval Philosophy* (London: Routledge, 1998), which emphasises continuity, and Eckhard Kessler, "Intellective Soul," in C. Schmitt et al. (eds.), *The Cambridge History of Renaissance Philosophy* (Cambridge: Cambridge University Press, 1988), emphasising reform.

[2] For commentaries and their frequencies, see the great catalogue of Charles H. Lohr, "Renaissance Latin Aristotle commentaries," in *Studies in the Renaissance* 21 (1974) and *Renaissance Quarterly* 28–33 (1975–1980), and 35 (1982).

Commentarii Collegii Conimbricensis Societatis Jesu in tres libros De anima Aristotelis Stagiritae.[3]

In that period several comprehensive Aristotle commentaries appeared, but the Coimbra commentary series was the most detailed of them. It was compiled by the Collegium Conimbricense, that is, the Jesuit philosophers working at Coimbra University in Portugal. Between 1592 and 1606, they published eight volumes of commentaries. These (with Fonseca's *Metaphysics* commentary) were the "advanced Aristotle course" of Jesuits. They were distributed internationally and reprinted many times. *De anima* was first published in 1595 and its last edition seems to have been in 1627.[4]

Francisco Suárez (1548–1617) is of course a famous figure, often called the last major scholastic and the first major modern philosopher. His *De anima* was printed only posthumously, in 1621, and was then slightly edited. It is important to notice, however, that the work stems from a much earlier time. It is based on his *De anima* lectures given in the 1570s and so belongs to the sixteenth century context, even if it stands out to some extent from the works of his contemporaries.[5]

The two works are closely interconnected since they represent the same Spanish Jesuit tradition, and Toletus in particular is a central figure in both of them.[6] Moreover, Suárez's *De anima* lectures must have directly influenced the Coimbrans, who were his personal colleagues. (Suárez worked for a number of years at Coimbra.) And again, the later, revised beginning of Suárez's book includes a few references to the Coimbrans.

3.2 The Role of *De anima I.1*

As we well know, Aristotle's *On the Soul* is an infuriatingly concise and strangely heterogenous work. Nowadays, philosophers concentrate on the second, and especially on the third, book – and the scholastics shared the same preferences.

---

[3] The editions used here are: Francisco Suárez, *Commentaria una cum quaestionibus in libros Aristotelis De anima*, ed. Salvador Castellote Cubells, tomo 1, Ediciones críticas de obras filosóficas 1 (Madrid: Labor, 1978); and *Commentarii Collegii Conimbricensis Societatis Jesu in tres libros De anima Aristotelis Stagiritae*, Lyon, 1600.

[4] Scholars believe that the *De anima* commentary was written by Emmanuel de Goes (1542–1597), who was in charge of the earlier commentary volumes, but it may contain fruits of genuine group work.

[5] Slightly before his death, Suárez began to rewrite *De anima* in a more concise form. He had written an enormous commentary on *Summa theologiae* and also wanted to show how all natural sciences could be fitted there to supplement and benefit theology. Of *De anima*, however, he got only the new beginning finished, and thus the published version replaced the introduction and the first disputation with the new text, but the rest remained nearly intact. A critical edition was published only in 1978–1991; it includes even the rewritten first part as an appendix.

[6] Concerning Jesuit anthropology, see Mabel Lundberg, *Jesuitische Anthropologie und Erziehungslehre in der Frühzeit des Ordens (ca. 1540–ca. 1650)*, Studia Doctrinae Christianae Upsaliensia 6 (Uppsala: Almqvist & Wiksell, 1966).

The first book consists mainly of not too friendly comments on Aristotle's prede-
cessors (Chapters 2–5). But the very first chapter of Book I (402a–403b) is dif-
ferent. Here, Aristotle examines the science of psychology in general, explaining
the nature, importance and function of these studies at the outset. Hence this chap-
ter is also interesting as one of his few concrete applications of his philosophy of
science.

Evidently, medieval authors saw Chapter I.1 as playing a special role by serving
as a general introduction to the *scientia de anima*. They discuss it as a prologue, and
occasionally even point to the tasks of every prologue. Thomas Aquinas explains
the idea very nicely: "Whoever writes a prologue has three things in mind: first, to
render us favourably disposed; second, to make us ready to learn; third, to make
us attentive. He renders us favourably disposed by showing the usefulness of the
knowledge; ready to learn, by setting out in advance the treatise's order and organi-
zation; attentive, by showing the treatise's difficulty." (*In De anima* I.1, 26–31.)[7]

These are indeed the themes that the commentators discussed: the worth and
usefulness of this study, its systematic organisation, its difficulty, and the problems
in defining it. As the scholastics saw it, Chapter I.1 sheds light on these preliminary
questions, and they try to explicate its full importance in their commentaries. Early
commentators, however, do not handle this chapter in any exceptional manner, but
discuss it in the same style as all the other chapters. Thus Albert, for instance, seeing
no particular significance, explains it. Thomas gives a clever commentary, but does
not distinguish the chapter from others. Equally, some Renaissance commentators,
such as Nifo, progress in extreme detail through the whole text.[8]

Others, such as Suárez and the Coimbrans, are more selective. They do not care
about Chapters I.2–5,[9] but they do discuss I.1, though in different ways. Suárez
begins his work with a long introduction, *Prooemium*, concerning the nature of this
science *de anima*, where he discusses the problems of this chapter without textual
references. The Coimbrans provide a short proem and then turn to I.1, commenting
on its details and settling various *quaestiones*.

The difference follows from the structures of the two books. The Coimbrans still
write a classical detailed commentary, giving the full Greek text and full translation,
and debate every conceivable difficulty phrase by phrase. They then turn to discuss
a few of their own questions in each chapter. Suárez's work is only nominally a
commentary: it is, undoubtedly, about the contents of *De anima*, but he writes a com-
pletely independent treatise, wholly neglecting Aristotle's disposition and selecting

---

[7] Thomas Aquinas, *A Commentary on Aristotle's* De anima, trans. Robert Pasnau (New Haven:
Yale University Press, 1999).

[8] Albert the Great (Albertus Magnus), *De anima*, ed. Clemens Stroick (Aschendorff, 1968);
Agostino Nifo, *Expositio subtilissima in tres libros Aristotelis De anima* (Venice, 1559).

[9] Suárez almost skips this material, giving only a condensed list of various old opinions about what
a soul is. The Coimbrans provide the text and add a brief classificatory sketch (*In DA,* 20–23),
saying that no closer analysis is needed because nobody supports the opinions criticised there any
longer.

his subjects freely. In his representation of Aristotelianism, Suárez approaches what is known as modern philosophy and its new style of shaping ideas.

3.3 Can there be any Science of the Mind?

If we progress as strictly and pedantically as possible, the very first question will obviously be: *Can* there be any science of the soul at all? Most authors fail to discuss such a scruple. After all, the matter was evident to Aristotle, and thus Thomas, for instance, immediately begins to use the word *scientia* for psychological knowledge. Similarly, in the later Aristotelian tradition it is generally taken as an obvious fact that there indeed is a science *de anima* in the class of sciences. However, some commentators raise a problem here. John of Janduno (d. 1328) already lists some imaginary objections, such as the following: (1) the soul is not intelligible because it is not even perceivable; (2) the understanding intellect must have a different nature from what is understood, hence the intellect itself cannot be understood; (3) understanding the soul would require that the soul moves itself which is always impossible. He refutes the objections as misunderstandings and offers a very illuminating argument to explain how knowledge of the soul is possible. As he points out, what is crucial is that the knowledge of the soul actually consists of the knowledge of its qualities, parts, functions and operations, and that *such* knowledge concerning the soul is quite possible, and is ultimately derived from knowledge concerning observables.[10]

This somewhat non-standard question occasionally appeared again later. Thus, Silvester Ferrariensis (d. 1526) lists practically the same objections as Janduno. He also rejects them in a rather similar way, pointing to the fact that "we have knowledge of [such] a subject whose predicates are known." Thus we have knowledge of the soul, too, but he complicates the matter by adding an enquiry about the methodological roles of "parts" and "wholes": that is, how can knowledge of various *parts* of the soul be said to constitute knowledge of the *soul* itself, in its entirety?[11] Toletus, in turn, treats the issue in an easier manner: "We are not discussing this because there is any controversy among Peripatetics (for everybody admits that there is a science of the soul), but to make the matter more explicit by means of a disputation" (1ra). He does not go into metaphysical details, but simply demonstrates that empirical investigation gives a lot of examples of the science of the soul. Indeed, the soul can be known *per definitionem*, because we obviously have affections and passions.[12]

---

[10] John of Janduno, *Super libros Aristotelis De anima quaestiones*, Venice, 1587, 9. Janduno's discussion is intelligent because the supposed objections are indeed imaginable misinterpretations of the Aristotelian doctrine.

[11] Franciscus Silvester Ferrariensis, *Quaestiones eruditissimae in tres libros De anima Aristotelis Stagiritae* (Rome, 1577), 2a–3a. He refers to the discourse in *An. post.* I.

[12] Franciscus Toletus, *Commentaria una cum quaestionibus in tres libros Aristotelis De anima* (Cologne, 1615), 1va–2rb.

The Coimbrans see no problem here. Their attitude seems to be close to that of Toletus: they speak synonymously of "the science of the soul" and "contemplation of the soul", and are happy to find a well-defined object of enquiry. Suárez also begins by referring to "this science", but adds a brief argument, which resembles that of Ferrariensis (*Prooemium*, n. 5). Suárez's argument is difficult to follow, but apparently his idea is to claim that there exists a science of animated, living things, which necessarily requires, as a part of it, this science of the soul. This is the first sign of his constant tendency to integrate this science into the larger context of other biosciences.

3.4 The Subject of Psychology

Supposing a science of psychology exists, the next problem will be to define what it is *about*: what is the *subject* of the science of the soul? Such a question can sound rather redundant. But the "subject of a science" is a technical notion belonging to the architecture of *Analytica posteriora*. And if we bear in mind the doctrine of Aristotelian scholasticism that the *soul* is the *form* of the *body*, then the question has obvious importance; it must be resolved before further progress. At least two main opinions seem reasonable in this framework: either the subject is the soul or the animated, living, body. In other words, it is either the form or the thing having the form. Classical scholastics were unanimous in that the subject must be the *soul*, although some discussion did apparently take place; for example, Janduno refutes a few counterexamples drawn from the theory of science.[13]

In the sixteenth century, many commentators still considered the matter more or less self-evident. Thus, Cajetanus and Nifo simply state that the subject is the soul, and make no comments. Other authors, however, see a problem here. For example, Javelli (d. 1539) begins with a whole question for this matter: "Is the soul the subject of the books *De anima*?" He comes to the conclusion that this is demonstrably the case, although "some people" have argued for the opposite. The problem remained alive: Francesco Piccolomini offers voluble argumentation about it, and finally supports a mediating solution where the soul is the "immediate subject" of the science, although the animated body is its indirect subject. (This probably approaches the traditional view.)[14]

Toletus, then, in his *Comm. in DA* q. 4, formulates the question clearly for his Jesuit followers. According to him, the first possible position would be that the subject of this science – "the subject, or object, also called the matter of the science" – is double: both the animated body and its principle or form, the soul. This, he says, was suggested by Paul of Venice. The second position, that of Apollinaris,

---

[13] Janduno, *Quaest. DA*, 12–13.

[14] Chrysostomus Javelli, *Quaestiones naturales super tres libros Aristotelis De anima*, Opera I, (Lyon, 1580); Franciscus Piccolomini, *Commentarium in tres libros Aristotelis De anima* (Frankfurt, 1602).

would be that the object is simply the animated body. And thirdly, it could be as Albert, Thomas, Giles, Janduno and "others universally say": that the object is the soul. Toletus at once accepts this third view. He supports it with an elementary argument, which, however, contains one interesting point (6va): he reminds us that a science must *define* its object, and this science defines only the soul and not any body.

The Coimbrans already – and only – discuss "the subject of these books" in their proem. They, too, say that Paul of Venice and "some recent philosophers" have supposed that the subject is the animated body, because this science is a part of physiology, the study of living beings. Most Peripatetics agree, however, that the subject is the soul, and this is the correct opinion. The subject of *Parva naturalia*, as a whole, is the animated body, but the preceding *De anima*, in particular, has a higher subject: the soul alone. It is fully possible that the subject of a part of the investigation differs from the subject of the complete investigation.[15]

Suárez begins his entire work with a detailed discourse in which he seeks simultaneously to find the proper object of this science and its place among sciences. He again points out that "although the object of a whole [i.e., of a whole investigation] is some totality, the object of a part can be some part" (*Prooemium*, n. 6). Thus he is strongly committed to the idea that psychology must be a part of a larger science of living beings. He even concludes (n. 7) that the opinions of Paul of Venice and of classical Aristotelians can be reconciled because of this. He apparently does *not* regard the controversial matter as a very prominent issue at all, because *De anima*, according to him, does not form a separate science *per se*, but necessarily constitutes a part of a more comprehensive whole of scientific knowledge (n. 5). But then, the proper object of *De anima* is really "the formal principle of an animated being, that is, the soul." In this way Suárez follows the classical opinion, but offers a surprising interpretation of it.

The remarks of these commentators have an interesting historical connection. An important debate had taken place earlier in Italy about the theory of the soul, when two unorthodox opinions became popular: Averroism and Alexandrism.[16] Nowadays it is admitted that these labels, in fact, define no clear-cut schools and have little to do with the historical Averroës and Alexander of Aphrodisias.[17] Such currents did exist in the fifteenth century, however. Paul of Venice, whom we now remember as a major logician, was one of the new Averroists, and Apollinaris, whom Toletus mentions, must be Apollinaris Offredi.[18] What Toletus says of him can be a crude

---

[15] The Coimbrans also emphasise, as does Toletus, the fact that *De anima* defines the soul, but it is difficult to see how their reference to *An. post.* I. 1 and 9 is relevant here.

[16] For a look at these issues in Paduan philosophy, see Section 2.2.3. in this volume.

[17] See, for instance, Paul Oskar Kristeller, "Paduan Averroism and Alexandrism in the Light of Recent Studies," in P.O. Kristeller, *Renaissance Thought* II (New York: Harper & Row, 1965).

[18] Apollinaris, *Expositio et quaestiones in libros Aristotelis de anima* (Venice, 1496). I have not seen this book. For more on Paul of Venice and his *Commentum de anima*, see Zdzislaw Kuksewicz, "Paul de Venise et sa théorie de l'âme," in L. Olivieri (ed.), *Aristotelismo veneto e scienza moderna* I (Padua: Antenore, 1983). Paul's opinion receives a fair treatment there (303–11).

representation of the Alexandrist views, later elaborated best by Pomponazzi, who does not appear in these Jesuit sources. (Cf. Alexander, *De anima* 24.18–25: "The soul is a form which comes into being when the bodies that are its substrate are combined and blended in a certain way [...]")[19] Well aware of this earlier controversy, the commentators respond to its themes. Thus, the Jesuits reject the Averroist and Alexandrist views in favour of the traditional theory, though from Suárez it receives a particular naturalist reading.

One special point remains to be settled. In those days, the question of *entelechy* came to the fore. The background is, simply, that Aristotle generally uses *energeia* as the pair of his *dynamis*, but in some contexts he uses the word *entelecheia* in its stead. In the crucial passage in *De anima* II.1, he says that the soul is "the first entelechy of the potentially living natural body" (412a27).[20] The classical scholastics had translated both *energeia* and *entelecheia* with *actus*, but in the eyes of some authors, the appearance of Greek texts and commentaries raised a problem here. Was entelechy a quite peculiar actuality in some way? Later, it became an important notion for certain early modern philosophers, such as Leibniz, but already in the sixteenth century it apparently generated speculation.[21] Thus we can read how John Eck insists that the subject of investigation, the soul, is a form and not any inherent vital principle, as a few authors had presumed. (As examples, he mentions George of Brussels and Bartholomeus Usingen.)[22]

To sum up, for the mainline authors the subject of psychology must be the *soul*. (Therefore I am not quite satisfied with the definition given in some neo-Thomist textbooks that the subject of psychology is *man*.[23]) And the soul is to be seen as a form informing the living body. Here, the body could in principle be whatever material thing that is living because of its informing form.[24] In other words, plants and animals also belong here. (Cf. Suárez, *In DA,* disp. 2.) But perhaps we may use less orthodox terminology and say that the subject of *psychology* – in contrast to *scientia de anima* – is merely the *human soul*.

---

[19] Alexander Aphrodisias, *The* De anima *of Alexander of Aphrodisias*, translated and commented by Athanasios P. Fotinis (Washington, D.C.: University Press of America, 1979).

[20] Concerning *entelecheia*, see Josef Stallmach, *Dynamis und Energeia* (Meisenheim am Glan: Anton Hain, 1959), excursus, 182–93. Stallmach begins his Exkurs thus: "Die Entelechiefrage ist oft genug aufgenommen und bisweilen auch resigniert liegengelassen worden".

[21] Stallmach's judgement is that *entelecheia* is mainly synonymous to *energeia*, in which case nothing could be concluded from the occurrence of this term.

[22] John Eck (Johannes Eckius), *Aristotelis Stagyritae philosophi De anima libri tres adiectis Eckii Commentariis* (Augsburg, 1520).

[23] Cf. R.E. Bourne, *Thomistic Psychology* (New York: Macmillan, 1957), 50: "Its proper subject matter is the human organism, or man, who by his nature, or essentially, is an intellectual organism, but who by his powers or virtually, is also a sensitive and vegetative organism."

[24] Now, of course, it is easy to see that this leads to great and far-fetching problems in defining *life*. Some discussion about that can be found in Dennis Des Chenes, *Life's Form: Late Aristotelian Conceptions of the Soul* (Ithaca: Cornell University Press, 2000), Ch. 3.

3.5 The Usefulness of Psychology

A preface, as we learned, must show the *usefulness* of the science. And indeed, Aristotle does so in I.1, albeit very briefly and obscurely (402a4–7). Therefore, it was natural that this issue gave rise to much explanation and development.

In medieval Aristotelianism, the first great author, Albert, emphasises that the science of the soul is highly useful because it provides highly important principles for all other fields of knowledge.[25] This is very close to Aristotle's own words: "[it] contributes greatly to the advance of truth in general." Similarly, Thomas says (*In DA* I.1.113–131) that in order to acquire full knowledge and mastery of the multiple fields of knowledge, one must first understand the relevant powers of the soul. Giles of Rome gives the often-quoted standard formula, which is an almost exact translation from Aristotle: "*animae cognitio plurimum conducit ad omnes alias scientias*," i.e. knowledge about the soul is greatly conducive to all other sciences.[26] Thomas and Giles continue by listing metaphysics, moral sciences and natural sciences, all of which benefit from psychology.

Though these early remarks were not very informative, later philosophers attempted to explain more clearly *how* psychology contributes to knowledge in other sciences. Buridan says very well that "the science of the soul is about the intellect and its operations, about sense and sensibilia and phantasms, and these are causes for the truth of cognition; therefore it concerns all truth that can be known." "An intellect that is ignorant of itself cannot be trustworthy about other things."[27]

Some Renaissance philosophers take the usefulness for granted as unproblematic,[28] but there also arose much discussion about it. Javelli, for instance, offers a complex, partly philological commentary (*Quaest. DA,* q. 4). He begins with Avicenna's note about different kinds of utility. In this manner, he can answer to an Averroist problem of Janduno: it is indeed possible that psychology is useful even for knowing first principles, because psychological knowledge then acts as a "productive", not a "formal" cause of insights. In other words, knowledge about the soul can assist and guide the philosopher in successful research quite generally. As such a universal tool, it is applicable to all sciences.

The theme of the usefulness of the science *de anima* was constantly repeated during the sixteenth century. In fact, it seems to have been especially cherished in popular textbooks. Later, Zabarella provides a full discussion with various digressions; even he is convinced that psychology is highly useful in several ways, both

---

[25] Albert, *DA* I.1.2. In fact he also suggests, quoting Cicero, that all very worthy things are bound to be useful.

[26] Giles of Rome (Aegidius Romanus), *Commentationes physicae et metaphysicae in De anima libros tres* (Oberursel, 1604), 559.

[27] John Buridan, *La Traité de l'âme [De prima lectura]*, ed. Benoît Patar (Louvain-la-Neuve: Éditions de l'Institut supérieur de philosophie, 1991), 6–7, cf. 179.

[28] Eck, *Comm. DA*, 4ra.

because it shows the principles of knowing and because it completes the structure of natural science.[29]

Toletus draws an interesting summary of the issue, a summary that may have been influential (*Comm. in DA*, 9v–10r). He agrees that psychology is useful for all sciences, but points out that it is especially useful for all knowledge of living things. (This alludes to the close connection between *De anima* and *Parva naturalia*.) Basically, psychology is useful for all sciences "because it deals with the effective principle of other sciences"; after all, "the soul is what abstracts, and so makes the objects of other sciences" (10r). Old philosophy has often been said to include no epistemology, in a modern sense; here Toletus seems to imply that psychology could fill its place by explaining the conditions of knowing.

Suárez and the Coimbra commentators both discuss the usefulness in detail, though amusingly, the Coimbrans begin with it whereas Suárez concludes with it. Suárez has a discourse about useful things in general and, more closely, about the usefulness of sciences (n. 31–32). According to him, this science is highly useful (*utilissima*), for it explains the soul and the intellect, and shows that many dogmas are completely reasonable: for instance, the dogma of immortality.[30] Moreover, because the soul contains the faculty of knowing, the science of the soul naturally benefits all sciences in their use of this faculty. Thus, dialectics and logic benefit from knowledge about the acts of the intellect, and moral sciences benefit from knowledge about what the soul strives for and how. The natural sciences enjoy a similar contribution, as do metaphysics and theology even.

For the Coimbrans, utility is paramount: psychology must be studied because it is so useful. After a note about divine phenomena, they explain that the science of the soul promotes moral sciences by displaying the mechanism of acts and the nature of virtues and vices. It also helps metaphysics by informing us of the nature of intelligible substances – even, analogically, of God. Moreover, it benefits other sciences by describing the nature of knowledge (although in the Coimbra text this is only a cursory remark). Finally, the Coimbrans present a long list of miscellaneous authorities praising the usefulness of psychology.

3.6 Is Psychology a Natural Science?

In the order of discussion of these authors, the next question very often is: Does psychology (*scientia de anima*) belong to *philosophia naturalis*, among natural sciences? The precise meaning of this question is perhaps not completely certain. In general, natural sciences are those which study only natural things and proceed by

---

[29] For Zabarella's views on "instruments of science," see Heikki Mikkeli, *An Aristotelian Response to Renaissance Humanism* (Helsinki: Suomen Historiallinen Seura, 1992). Cf. also Piccolomini, *In DA,* 409–12.

[30] Of course Suárez is responding here to famous earlier controversies. On the other hand, this is also one of the many points in this material which forebode the texts of Descartes.

means of the ordinary, essentially empirical, cognitive faculties. The way in which this notion was understood may, however, have undergone changes and refinements.

Albert stated very strongly that psychology is one of the natural sciences, for these sciences study bodies, even moving bodies (*corpus mobile*), and the soul must be seen as an essential principle in such cases; even in man, the soul is a perfection of the body. Similarly, Giles of Rome asserts: *anima est praestantissima pars corporis animati*, the soul is the most prominent part of the animated body.[31] Thomas himself does not discuss this matter, and apparently did not find it central to his thought.

On the other hand, Janduno asks the question "Is the science of the soul natural?" (q. 2). His answer is affirmative. This is clear for the vegetative and sensitive soul, but knowledge about the intellective soul also belongs to natural science in so far as it pertains to the soul's operations, faculties, etc. Among nominalists, Buridan speaks without further explanation of *scientia de anima* and "other natural sciences".

Gradually, the question became more troublesome because of growing dualistic tendencies. In 1520, Eck attempts to draw a systematic picture of the matter. According to him, consideration of the soul is of two kinds: *in se et suapte natura* it is metaphysics, but *a posteriori* it is natural science (2ra). Such an answer, which seems to have been already well-established at the time, implies that the soul *as an entity* belongs to the field of metaphysics, whereas *mental principles* belong to natural science. Eck adds that some commentators had excluded the most intellective principles of the soul from physiology (Themistius, Simplicius, Nifo), but Alexander, Avicenna and Averroës, on the other hand, advocated its position among the natural sciences.[32]

The issue generated wide debate, and one easy proposal at the time was that psychology could lie somewhere between the natural sciences and metaphysics.[33] But the old Aristotelian answer also had stern defenders. Javelli, for instance, steadfastly maintains that this science is natural, and provides a cunning answer to a methodological objection: the subject of one part (psychology) of a natural science need not literally be a part of the subject (animated body) of the whole science, but can be a passion or principle of it (*Quaest. DA*, 614); thus it is not impossible that the science of the soul is natural.

In fact, the problem has great importance, which is perhaps not immediately obvious. The decision here will affect the whole conception of the mind. Conscious of

---

[31] Albert, *DA*, 1–2, Giles of Rome, *Comm. in DA*, 558.

[32] Eck also refers to the beginning of *De sensu et sensibilibus*, where "Aristotle seems to count *scientia de anima* among physical fields."

[33] See Lucillus Philaltheus, *In libros tres Aristotelis De anima Commentaria absolutissima* (Turin, 1579), 2. Themistius and Simplicius placed the knowledge of the foundations of intellect outside the natural science, to a status resembling that of mathematics. Psychology is therefore a "*scientia media*" between the natural and metaphysical spheres. Nifo attempted to defend this seemingly unpromising position by means of the double nature of the soul. See Edward P. Mahoney, "Agostino Nifo (ca. 1470–1538) on the *scientia de anima* as a 'mathematical' or 'middle' science," in R. Työrinoja et al. (eds.), *Knowledge and the Sciences in Medieval Philosophy III* (Helsinki: Luther–Agricola Society, 1990).

this importance, Cajetanus examines the problem of the naturalness of psychology in detail (n. 56–72, p. 49–58).[34] He asks whether the soul belongs wholly to the field of natural philosophy. *De anima* is situated in that context, but contrary arguments exist as well. For instance, placing the soul wholly within the field of natural philosophy would apparently make all science natural; moreover, the intellect seems not to belong to physics since it involves no motion principle and has no final causes, as do the physical principles.

The solution of Cajetanus is to view the soul in two ways. On the one hand, physics never abstracts from sensible matter. Metaphysics, on the other hand, does.[35] Now, which is the case in psychology? There we can study the intellect either absolutely or in some mode. Absolutely taken, it has nothing to do with matter and therefore belongs to metaphysics, but in its functions, that is, "taken together or conjoined with phantasms, it concerns sensible matter." Here we have *two ways of thinking* of the intellective soul: either absolutely or with regard to the body (n. 59).[36] The two ways lead to metaphysics and to natural science, respectively.

Cajetanus refutes a number of supposed counterarguments and finally comes to the basic question: Does the human soul *quidditatively* – according to what it is – belong to natural philosophy? Immortality, for instance, is a strong argument against its naturality, but Cajetanus wants to take a middle way (n. 70) and states that the difference lies in whether we are interested in the entity itself or in its acts. Making one's choice here leads to two different enquiries.

Cajetanus was hardly the first advocate of such a "double view", but this view became very common after him. The Coimbrans, for example, remained inclined to accept it. They devoted the only *quaestio* in their proem to the issue of the naturalness of psychology. There they first (article 1) listed various opinions that place the study of the intellective soul either in metaphysics, in physics, or in an intermediate position. In their answer (a. 2), they refute the intermediate view and state that the soul can be seen in different ways. Seen wholly in itself, transcending matter, it concerns metaphysics, but with regard to bodily functions, it belongs to natural science. Moreover, as regards its own nature and essence, the soul belongs to natural science because it is *essentially*, by its very nature, the form of the body. This is an interesting point that further underscores the problem that occupied Cajetanus and subsequently Suárez.[37]

---

[34] Thomas Cajetanus, *Commentaria in De anima Aristotelis*, vol. I, ed. P.I. Coquelle (Rome, 1938).

[35] See *ibid.* n. 58; Aristotle, *Met.* VI.1, 1026a5–6. It is easy to see that the sensitive and vegetative soul, busy with sensible matter, belong to natural science; a problem arises only with the intellective soul.

[36] Cajetanus states that the soul and the body are "partly conjoined and partly separate", referring to *Phys.* II.2, 194b12–14. But what actually can this "partly" mean?

[37] In *De part. an.* I.1 (641a33–b10) Aristotle clearly seems to place the rational soul in metaphysics. The problems of Cajetanus stem ultimately from this passage. The Coimbrans conclude (a. 3) with a somewhat obscure answer to it, and Suárez attempts to deal with the same passage in his *Prooemium*, n. 20–23. According to him, Aristotle speaks there about the intellect "absolutely and in itself", and in particular wants to stress that celestial souls are not natural.

Suárez, who does not usually like Cajetanus, largely shares the same opinion in his *De anima*, though he gives it a clearer ontological form. He asks: Does the *scientia de viventibus* belong to "physics"?[38] (Note that he does not say *scientia de anima*.) He takes the opposing arguments very seriously. Firstly, the necessary abstraction from all material being leads from the natural sciences to metaphysics. Secondly, living things as such (*viventia ut sic*) are not natural, because their defining feature, soul, need not be natural. Thirdly, the rational soul is purely immaterial.[39] Now, to evaluate these points, what is "living" is that which contains the principle of operations promoting its perfections, and that can be fully material and belong to the scope of physics. Thus it is correct to conclude that purely natural – that is, irrational – operations surely belong to physics (cf. *De part. an.* I.1). And purely immaterial operations, such as those of the angels, instead lie outside the field of physics. In this way, only the human rational soul presents a problem. On the one hand, it is an independent spiritual entity, *entitas spiritualis independens*, yet on the other hand, it is the form of the body, *forma corporis*. In the former sense, it belongs to metaphysics; in the latter sense, to natural science, or "physics".

In a way, this is in accordance with what Suárez said about the subject of psychology, but he tends to "ontologise" the duality of the two studies. Unlike Cajetanus and the Coimbrans, he does not refer to our perspectives or interests at all. This at once leads to the crucial question: How can the soul, the same thing, be *both* an independent entity and simultaneously a form of another, bodily thing? How could this even be conceived in the framework of Aristotelian metaphysics? Apparently, what animates a rational being would now be another entity, something different from what is animated, or at least this would be the case "metaphysically seen".[40] The same crucial problem threatens all the Aristotelians who wish to emphasise the substantial nature of the soul. Suárez, like others, must encounter the dilemma that Aristotelian metaphysics and strict dualism are basically incompatible.

Curiously, Suárez does not dwell on this discrepancy even in his metaphysics. Instead, in a noteworthy passage (*Disp. met.* disp. 1 s. 2 n. 19–20) he asserts that the soul, though a subsistent entity, is essentially related to matter and is known only as such. The outcome is that "consideration of the [rational] soul belongs to the last and most perfect part of natural philosophy." No distinct, purely metaphysical level is mentioned. This apparently indicates that metaphysical reasons led Suárez to abandon excessive dualism in favour of clearer Aristotelianism. Indeed he refers to Toletus, who had come to the conclusion that only the traditional opinion was correct (see Toletus, *Comm. in DA,* q. 2).[41]

---

[38] See Suárez, *In DA*, n. 10–2.

[39] This silent assumption is, of course, crucial for the whole dispute, but was always accepted as unquestionable.

[40] For more on these problems see Des Chenes, *Life's Form*, 76–9.

[41] The soul figures very little in the great *Disputationes metaphysicae* of 1597. See Francisco Suárez, *Disputationes metaphysicae* (Paris, 1866; repr. Hildesheim: Georg Olms, 1965). In fact, Suárez consciously excludes it (disp. 1 s. 2 n. 20), though admits that "separate souls" probably

Whether or not psychology is natural has no direct bearing on the more famous debate about the immortality of the human soul. This is not a subject for *De anima* I.1 at all, but perhaps something nevertheless ought to be said about it, as immortality was an important issue for everybody studying the soul. Roughly speaking, there were three main questions: (1) Is the individual human soul immortal? (2) Is this a fact that can be shown by natural means? (3) What was Aristotle's opinion here?

Pomponazzi's initiative made individual immortality an intensely debated problem, and many important works were related to that controversy.[42] If the rational soul is indeed demonstrably immortal, this poses a weighty point in the doctrine of substantial forms. Moreover, immortality obviously leads to the question of separate souls, that is, bodiless souls between death and resurrection. The issue became very popular, and must have reinforced dualistic tendencies.[43]

Finally, it deserves mention that the Coimbrans honestly admit that Aristotle never discussed this subject. Therefore, they append a special treatise on the separate soul, *Tractatus de anima separata*, by Balthasar Alvarez. According to him, immortality can be shown by natural light, as the Lateran Council recommended in 1513. First of all, a rational soul is a spiritual substance, *subsistens per se*; thus, it is immortal. The spirituality can be inferred from its operations, as it functions with abstract objects. The second main argument is that the soul cannot be corrupted without the generation of a new substance, but what could that substance be?[44] Aristotle would have to have been stupid had he not seen these reasons, and he most certainly was not. Alvarez, however, admits that the arguments are not "equally strict and evident as those of Euclid" (510).

---

belong to metaphysics. In his revised *De anima* he clearly returns to the orthodox opinion that *scientia de anima* is a natural science (see, for instance, Ch. 2).

[42] For more on the themes of the debate, see, for example, Martin L. Pike, *Pietro Pomponazzi: Radical Philosopher of the Renaissance* (Padua: Antenore, 1986), Ch. II; Étienne Gilson, "Cajétan et l'humanisme théologique," *Archives d'histoire doctrinale et littéraire du moyen âge* 30 (1955), 113–36; Kessler, "Intellective soul," 500–07; Des Chenes, *Life's Form*, 45–50; Kärkkäinen & Lagerlund, "Philosophical Psychology in 1500". Javelli wrote a fine short survey of the issue, *Tractatus de animae humanae indeficientia*, 1536. He favoured the traditional view that Aristotle surely regarded immortality as naturally demonstrable.

[43] In his *De anima* Suárez offers a whole disputation (disp. 14) about separate souls. Unfortunately, this discussion is beyond the scope of this chapter. In the revised edition he approaches orthodox Thomism. An overview of Suárez's doctrine of immortality appears in Salvador Castellote Cubells, *Die Anthropologie des Suarez: Beiträge zur spanischen Anthropologie des XVI. und XVII. Jahrhunderts* (Freiburg: Karl Alber, 1962), 196–201.

[44] Balthasar Alvarez, *Tractatus de anima separata*, disp. 1 a 3, appended in *Commentarii Collegii Conimbricensis Societatis Jesu in tres libros De anima Aristotelis Stagiritae*. The other six arguments are weaker. One of them alludes to the natural urge toward immortality. Thus, for instance, people build "mausoleums, colossuses, pyramids and columns" to immortalise themselves. Another argument stresses the moral consequences of the matter, as does the Preface of Descartes's *Meditations*.

3.7 The Order of Biosciences

The following question, admittedly, is possibly not such a burning one for us any more, but it did worry the professional Aristotelians. Supposing that psychology belongs among the natural sciences, what exactly is its proper place in their system? Here we must remember that the decent ordering and classification of sciences was a very dear common theme to Renaissance philosophers.[45] In the ordinary corpus of the Aristotelian works, *De anima* is the first of all natural works. Now the question is: Is this position correct, and are there any deeper arguments to support it? Yet, it is worth noticing that this discussion focuses on the *logical order*, the relations of mutual dependence between studies, which can then guide the correct *order of exposition*.

This, apparently, was an old topic: Janduno (*Quaest. DA*, 5) already mentioned some small disagreement, mainly about the place of the *Meteorology*. Much later, Eck remarked that different editions of the Aristotelian text used somewhat different orders. (Indeed, several new editions and translations of the texts were then available.)

After all, it sounds very plausible that the sciences of living nature, beginning with the most general one, follow general physics and the sciences of lifeless nature, but which of them is this, actually? The traditional view was that the first bioscience must be the science of the soul, *de anima* (cf. Albert, *DA*, 3a), but dissenting opinions arose, and there seems to have been disagreement about the correct order. Piccolomini quotes four possible views and comes to the conclusion that the traditional order is partly antiquated.[46]

Suárez discusses such views in n. 8. Some people, he says, believe that *De anima* should, on the contrary, be the last science of the animated nature. He refers to Nifo, who indeed appears to think that psychology precedes other biosciences only in nobility.[47] But for Suárez, "in distinct cognition we must proceed from parts to the cognition of the whole," and thus it is right to begin from the fundamental common part of all animated things, which is the subject of *De anima*.

Other writers had said that just *De partibus animalium* ought to precede *De anima* "because it explicates the organisation of the body by which the soul is defined." Suárez admits that this view is somewhat reasonable, since the soul and the animated body are "in cognition interdependent," but regards the familiar simple opinion as logically more satisfactory. In fact one gets the impression that this detail was not very important to him. When he relinquished the old commentary form, many old Aristotelian items also became matters of convenience.

---

[45] See Katharine Park, "The Organic Soul," in C. Schmitt et al. (eds.), *The Cambridge History of Renaissance Philosophy* (Cambridge: Cambridge University Press, 1988), about Gregor Reisch and others; also William A. Wallace, "Traditional Natural Philosophy," *ibid.*, 209–12.

[46] Piccolomini's book contains a long and confused, but interesting, discussion here. *Comm. in DA*, 386–91.

[47] Nifo, *Exp. in DA*, 3.

In the very beginning, the Coimbrans, on the other hand, offer an emphatic argument to show that the old "order of these books among other parts of physiology" is definitely the best. Alexander and Averroës had placed *De partibus animalium* before *De anima*, "because contemplation of matter precedes contemplation of form," but this is fallacious: the ease of results does not imply true priority. "The more generally valid and more common things must be treated before others," and hence the science *de anima* is the first bioscience. According to the Coimbrans, it is the first even in the sense that it does not need much assistance from the others.

3.8 The Worth of Psychology

The two preceding sections discussed the place of psychology in the classification and succession of sciences, but there also was a third ordering: it was customary to weigh the *worth* of different sciences. Why is a science good, and why are some sciences better than others? Sophisticated discussion here demands distinctions of goodness: what is good can be useful, delightful, or "honestly good".[48] Utility was already mentioned, and delightfulness is not relevant in this context, so the question concerns pure goodness, which can also be called honourableness or nobility. (It is important to remember that differences in the nobility of sciences tell nothing of their necessity.)

All sciences are honourable, but the problem arises because Aristotle gives two criteria for relative honourableness. "While knowledge of any kind is a thing to be honoured and prized, one kind of it may, *either* by reason of its greater exactness *or* of a higher dignity and greater wonderfulness of its objects, be more honourable and precious than another" (*DA* I.1, 402a1–4). It is unclear whether the two criteria must coincide, and this question caused much debate. Moreover, *akribeia* or exactness was usually translated as *certitudo*, which gives a bit different tone.

Thomas Aquinas (*In DA* I.1, 78–97) expressed a classical and widely supported solution. Indeed, some sciences are more certain or rigorous than others, and some concern things that are better. "Even so, that knowledge is better that concerns things that are better and more worthy of special honour. The reason for this, as the Philosopher says in the *De animalibus*, is that we would rather know a little about the highest things, the things more worthy of special honour, even if we know it topically (i.e. with probability), than know much, with certainty, about things less lofty in status."

This solution was universally known. With or without Thomas's special explanation, many philosophers, even those who were far from Thomas's thought, repeated his conclusion. It amounts to a decision that the worth of sciences should be judged primarily from a metaphysical standpoint, on the basis of their subjects;

---

[48] *Bonum honestum, delectabile et utile*. This was extracted from *Eth. Nic.* I.8 and was quite standard; see, for example, Thomas Aquinas, *In quattour libros Sententiarum* II, in *Opera Omnia* 1, edited by Roberto Busa (Stuttgart: Friedrich Frommann Verlag Günther Holzboog, 1980), dist. 41 q. 2 a. 2 exp. The surprising term "honest" means something like "proper" or "pure".

the methodological merit is a separate criterion, which is quite genuine, but less profound.[49] And thus, *scientia de anima* can be judged the most honourable natural science simply because of its noble subjects. At least for Thomas, this seems a well-considered opinion, but in fact he immediately continues that this science has *both* of the Aristotelian excellences: it is certain and concerns noble things. (As many authors note, these comparisons must be limited to natural sciences, and exclude theology and metaphysics.)

Most authors agreed that psychology is the noblest natural science *both* in regard to its certainty *and* to its objects. That, indeed, was Aristotle's own suggestion: "on both accounts we should naturally be led to place in the first rank the study of the soul" (*DA* I.1, 402a4). Obviously, this requires some argument to support the maximal certainty of psychology. Albert already stated that its demonstrations are sure because "nothing is more manifest than the source and origin of all cognition" (*DA,* 4a), and Giles simply said that it is very certain because it deals with "what we experience in ourselves" (*Comm. in DA*, 559). At least from Albert onwards, we thus find constant emphasis on the privileged first-person view. ("Nothing is more known than what one experiences in oneself." This is not from Descartes but from Ferrariensis [*Quaest. In DA,* 6b].)

Albert and Giles saw no problem with their position, but later the issue of certainty was much discussed. For instance, are not some obvious physical truths and demonstrations more evident than some results of psychology? One explanation could be a distinction between certainty for us and certainty *secundum naturam*: the natural certainties of psychology are not always apparent to us.[50] Furthermore, is psychology really "more certain and noble" than mathematics, as Averroës implies? There were several suggestions: some authors quote Burley's amusing idea that mathematics is a little more certain than psychology, but psychology concerns greatly nobler things, hence the sum total favours psychology. Psychology could also be considered highly certain simply because it is certain about such high things.[51]

A more interesting explanation for the mutual certainty of psychology and mathematics stems from Albert, who had suggested that the notions of knowledge and certainty are basically psychological (*DA*, 4a); later authors elaborated upon this idea.[52] Its point is to argue, using a slightly anachronistic formulation, that the certainty of psychology provides the *source* and *conditions* of the certainty of mathematical thinking. In this sense, the certainty of psychology precedes that of mathematics.

---

[49] Cf. *Met.* XI.7, 1064b1–6. The matter had already been troublesome for ancient commentators. See Giancarlo Movia, "Note sul primo capitolo del 'De anima' di Aristotele," in *Scritti in onore di Carlo Giacon* (Padua: Antenore, 1972).

[50] Buridan had disapproved of such an idea, saying that the full certainty of psychology concerns only what is actually experienced (*Tract. DA*, 6). In this way he admits that many results of the psychological science remain uncertain.

[51] Ferrariensis toys with this slightly sophistical view (*Quaest. in DA*, 6), referring to Linconiensis (Grosseteste), but Toletus strictly rejects it as irrelevant (*Comm. in DA*, 9va).

[52] For example, see Javelli, *Quaest. in DA*, 616a, Nifo, *Exp. in DA*, 6–8, 13.

However, one moderate opinion admitted the unquestionable certainty of mathematics, observing that mathematics simply does not belong to the natural sciences in this connection.[53]

The nobility of objects is easier to believe, and was often commented upon rather routinely. Some standard counterexamples were repeated, and even the way they were refuted was conventional. For example, the prime mover (*Phys.* VIII) and the eternal heavenly forms (*De caelo*) may seem nobler because of their immensity and eternity, but those perfections do not concern the aspects that science discusses.[54]

The order of nobility of the sciences does not seem to be an equally intense question for later authors. The Coimbrans provide a short commentary note (p. 12–3) in which they calmly expose Aristotle's text and say that the science of the soul must be *inter nobiliores*, which strictly speaking means only that it is *among* the noblest, *one* of them. Suárez studies the matter in greater detail. His conclusion (n. 28) is that *scientia de anima* "holds the third place among the sciences. This is seen from the nobility of its objects." In other words, it comes after theology and metaphysics, which hold the first two places in nobility.

Psychology occupies its high status because of the nobility of its object. Thus Suárez arrives at the well-established Thomist position. However, he considers it necessary to dwell on this point in order to comment on the old objections that the whole animated body or the prime mover could be nobler that the soul. He believes that Cajetanus had refuted these objections in a mistaken manner, and thus must solve them again.[55] Suárez's opinion is grounded on a rigorous division between sciences: in so far as the prime mover belongs to the natural sciences, it is to be seen only as a physical mover, and that is less noble than intellectual activity. Cajetanus, in contrast, sought to find some vital, psychological aspects in the prime mover.

3.9 Why is Psychology so Difficult?

As Thomas told us, by showing the difficulty of a treatise in the prologue, the author renders readers more attentive. This is surely a reasonable point with which everybody agreed. And it was also much emphasised because the science *de anima* is difficult in the greatest extent: it requires supreme effort and skill from the philosopher. Aristotle himself had pointed out immediately after the remark about the worth of this science: "To attain any knowledge about the soul is one of the most difficult things in the world" (*DA* I.1, 402a10).

Aristotle continues by pointing out a number of fundamental difficulties that research here must confront. These include ontological problems about the notion of the soul. His outline is rather unclear, and commentators developed it to a more

[53] See Cajetanus, *Comm. in DA,* nn. 3–9, pp. 18–21.

[54] E.g., Ferrariensis, *Quaest. in DA*, 8b-9a, Eck, *Comm. DA*, 5v, Toletus, *Comm. in DA*, 9va.

[55] See Suárez, *In DA*, Prooemium, n. 29; Cajetanus, *Comm. in DA*, n. 15–9, p. 24–7.

explicit form. Thus they produce lists of basic ontological or categorical questions,[56] but these problems are basic problems *within* the philosophy of the mind: they do not really concern the status of this enquiry itself. Seen in this way, that material, which comprises most of *De anima* I.1, does not pertain to the higher-level issue about the position of psychology in philosophy.

There is, however, one purely theoretical puzzle about the difficulty of psychology: If psychology is extremely certain, how can it also be so difficult? One could think that what is very certain must be obvious to everybody, and so the two characteristics of the science seem to conflict. This is a justified question, and Javelli's answer is interesting. He says (*Quaest. DA* q. 5) that the difficulty of knowing can derive from two causes: either the objects are especially abstract, potential, relative, etc., or they are hardly manageable for our intellect, which must always use its method of phantasms. Against Janduno, Javelli insists that the objects of psychology are quite actual, but in the case of the intellective soul, they fall outside the sphere of senses, and therefore "knowing these forms is difficult not because of the forms, but because of our intellect."

Suárez's concise answer (*In DA*, Prooemium, n. 33) is equally noteworthy. He admits that this science does enquire about evident things, but remarks that "in what we experience, there is much that can be enquired only with considerable difficulty." In other words, in these evident things there is much that is not evident at all. Clearly, we experience some psychological phenomena in ourselves, but what are the principles, potencies, and processes behind them? What are the laws of psychology? So, in one sense this science is "easy and very certain," yet in another sense, "difficult and obscure to us." Of course, the rational soul is the most difficult to understand, since its operations are most abstract. Here, Suárez implies that psychology ought to explore very difficult analyses for very obvious phenomena.

The issue of ease and difficulty rounds off the traditional main questions in the prologue of *scientia de anima*. In the sixteenth century, these questions were present in all commentaries, and the leading commentators formed their opinions while fully aware of the systematic importance of the matter. In other words, the meta-theoretic introduction was clearly seen as defining the place and role of psychology in the whole of philosophy; the field had reached a new level of self-consciousness.

All this was accomplished within the framework of the philosophy of old masters. But their works were considered insufficient and in need of constant reinterpretation and elaboration. In fact, they were not even used very much in everyday studies. I cannot do better than to quote Cranz: "In a long development the sixteenth-century Aristotelians had developed a whole new superstructure of terms and ideas through which to discuss Aristotle. These terms and ideas had made it possible to modernize Aristotle and in some cases to move to new non-Aristotelian or even anti-Aristotelian positions."[57] The notion of psychology is one such case where the

---

[56] Thomas already works in this way. For a fully articulate version, see, for instance, the list of ten questions in Cajetanus, *Comm. in DA*, p. 15.

[57] F. Edward Cranz, "The Publishing History of the Aristotle Commentaries of Thomas Aquinas," *Traditio* 34 (1978), 179.

new understanding of Aristotelianism seems to differ from that of the Middle Ages, evolving towards the pursuits of modern philosophy.

Nevertheless, the overall impression of these detailed discussions is admittedly somewhat chaotic. Scholars investigated and debated innumerable details, but it is hard to gather this wealth of material into a full and coherent view of the philosophy of psychological processes. In that sense, Suárez's *De anima* is something of an exception because of its well-organised disposition. This achievement was probably connected to the new format adopted in this work, which abandoned the traditional commentary pattern. Thus, it gave a promise of modernised Aristotelianism, a promise which, however, saw little development in the profoundly changing intellectual movements of the seventeenth century.

Chapter 4
Cartesian Psychology – Could There Be One?

Mikko Yrjönsuuri

Abstract The chapter examines what it would mean to talk about "psychology" in Descartes' terms and argues that within the Cartesian framework we cannot really formulate the questions that are posed by contemporary psychologists. This results from the fact that psychological topics can be found on all three levels of Cartesian science: in metaphysics, in physics and finally in the applied sciences, such as medicine and morals. The aim is to show that the sensory and vegetative functions are often taken together by Descartes. Therefore, the Cartesian system does not recognize any principal difference between sensory functions, such as vision, and vegetative functions, such as digestion. Humans can be conscious of both functions being operative in their bodies, but neither function presupposes the existence of a soul. This interpretation emphasizes the importance of Descartes' anatomical writings, including the manuscript *Primae cogitations circa generationem animalium*, which have been neglected by most contemporary commentators and scholars of Descartes.

As the university discipline called "psychology" nowadays understands itself, it was shaped in the late nineteenth century. It is a very wide-ranging subject, addressing a large variety of issues but usually not from a Cartesian perspective. Also, it is clear that when developing his own universal system of sciences in the second quarter of the seventeenth century René Descartes did not explicitly propose any such discipline. In his system, there is no evident place for a discipline that would directly correspond to what we now know as psychology. In such sense, it is clear that Descartes was not concerned with the problems of scientific psychology: there was no Cartesian psychology.

But could there be one? Could we, from the Cartesian perspective, demarcate a field of study that would appropriately be called psychology? How would it be defined within the Cartesian system of sciences? The main problem to be addressed

M. Yrjönsuuri
Department of Social Sciences and Philosophy, University of Jyväskylä

in answering such questions is of course the same as the one encountered in the history of philosophy more generally. We live in a different time from Descartes', and if he were introduced to what is now known as psychology, his philosophical stance would change. So in this situation what could count as a "Cartesian perspective"?

It is evidently to some extent a matter of taste how we put modern psychology within the framework of Cartesian sciences. But it is equally clear that there are more or less correct ways of doing this, and, even more importantly, there are more or less illuminating ways of doing this here and now. My guiding principle in the examination at hand is that history provides the most illuminating answers when it is done in a way that is historically as true as possible. Nevertheless, there is no single way of constructing such interpretations, nor is any true interpretation final, since all depends on what kinds of issues of our own time we are interested in and, thus, on how we put questions to historical texts. An interpretation is an answer to a question applied to a text, and, as questions develop, interpretations are bound to change as well.

It is possible to produce a historically true and complete representation of the positions held in the Cartesian corpus: by reprinting the texts. But translating a text already forces one to pose one's own questions to the text, thus, the translation has to be made with an understanding of how the texts relate to the intellectual issues contemporary to the translation, and, thus, the positions represented in the translation differ from those in the original text. *A fortiriori*, this is true of any discussion of the texts.

As we start to address the problem of a Cartesian psychology in this spirit, it soon turns out that we are forced to look closely at a variety of metaphysical and epistemological questions. Also, we cannot avoid the more practical but no less basic question of what is at issue in the sciences more generally. When we try to explicate what a Cartesian psychology would look like, it is necessary to examine Descartes' way of building the system of sciences. What would the role of psychology be in his system of sciences? Why would it be studied, and what kinds of methods and results could we expect in a Cartesian psychology?

4.1 Defining the Subject Matter of Psychology

Cartesian psychology cannot be a "science of the soul", at least not in the classical Aristotelian sense. In early seventeenth-century university education, courses with the title "On the Soul" had a traditionally well-established role. It seems appropriate enough to characterize this field as a university discipline in roughly the same manner as disciplines were characterized in late twentieth century universities. For the purposes of this chapter, we need not take a stance on whether this discipline should be identified as the same or even as closely corresponding to twentieth-century psychology, since it is clear that Descartes rejects the identification of this classical

discipline as a unified whole. In his system, the traditional courses on the soul would have to be radically reorganized.

The reason why Descartes rejects the traditional discipline is similar to the reason we hesitate to accept it as the medieval precursor of modern psychology: the Aristotelian discipline finds too many souls in the world. The Aristotelian "science of the soul" aims at a scientific explanation of how living corporeal things change and move and how they react to external influences. In this science, "soul" means whatever is the principle of life, explaining nutrition, growth, and a number of other functions assigned to organisms. Thus, even plants have souls.[1] However, insofar as we can admit that modern psychology is a science of the soul, we are using a concept of soul that makes it nonsensical to claim that plants have souls. The scientific explanations of how plants grow and interact with each other are not and cannot be psychological. Plants have no psyche in the relevant sense, and they are not studied in psychology.

Descartes too rejects the plant psyche. In his view, everything that happens in plants can be explained mechanically without reference to anything like a soul. In a further rejection of the Aristotelian picture, he also rejects that any animal outside the human race has a soul. To be more exact, he rejects that plants and animals have souls like those of humans, and he prefers to reserve the word "soul" (Latin *anima* and French *âme*) for the kind of thing we humans are conscious of having and for which the Church claims immortality. A science studying the soul would thus have to be limited to studying humans. This approach differs greatly from the Aristotelian understanding of the subject matter dealt with in the science of the soul.

In the classical Aristotelian picture, powers of the soul can be classified on three levels, often called the vegetative, sensitive, and intellectual levels. In his unpublished writings, Descartes occasionally operates with this distinction, and in those contexts he explicitly opts to say that only the intellectual level should be understood in terms of the soul. While we, on the basis of the twentieth-century tradition, would rather use a concept of soul that leaves only the vegetative level out, Descartes locates the sensitive functions at the corporeal, soulless level in all texts where he is confronted with this three-level model. Indeed, his way of drawing the line between the soul and the body differs not only from Aristotle but also from any corresponding discrepancy in use nowadays. In particular, I wish to show here that his distinction cannot be used to yield anything like our distinction between psychological and physiological phenomena, since some crucial phenomena that Descartes attributes to the body clearly fall into the scope of the discipline of psychology as it is studied today.

Let us look more closely at some relevant texts. First, Descartes wrote a letter to Regius (Henri Le Roy; 1598–1679) in May 1641. At the time, he was composing replies to the objections that were to be published with the *Meditations on First Philosophy*. In the letter, he aims at making clear how living entities differ from lifeless ones and how the word "soul" ought to be used in this context. It seems that

---

[1] See, e.g., Aristotle, *De anima* book I, Ch. 1; book II, Ch. 1; and book II, Ch. 4.

Regius had employed distinctions between (1) the power of growth, (2) the power of movement, and (3) the mind (CSM 3:181–82; AT 3:371–72).[2] The three members of the distinctions correspond to the three Aristotelian levels of the soul, and indeed Regius appears to prefer using the word "soul" for them. Descartes disagrees, saying that, in his view, the two first mentioned powers are of the same kind as each other but totally different from the mind. He admits that the two powers are called "souls" in plants and in non-human animals, but he wants to make it clear that the human soul is a soul in a very different sense. Even more interestingly, he concedes that these powers "exist also in human beings; but in the case of human beings they should not be called souls." As Descartes sees it, these powers are not related to what is the "first principle of action" in humans, thus implying that the concept of the soul is to be spelled out in terms of what is the first principle of action. However, the admission that these powers exist also in humans implies that the functions to be explained by these two powers, related to the Aristotelian plant and animal souls, are to be found in humans as well. Human growth is to be explained through the same principles as plant growth, and the same goes for the functions traditionally explained by the sensitive soul.

For a more thoroughgoing analysis of what Descartes would assign to these categories, we have to turn to an earlier text: the *Treatise on Man*, presumably written in the early 1630s. There, Descartes formulates his scheme with the vigor of a young, innovative thinker. At the beginning of the text, he adopts the strategy of describing first the body, then the soul, and thirdly "how these two natures would have to be joined and united in order to constitute men who resemble us." In fact, he withdrew from publishing the work, and indeed it seems that he never completed it (although in a letter to Mersenne 5.4.1632 he claims that it is practically finished [CSM 3:36; AT 1:242], and the *Discourse on the Method* refers to a part that we do not have [CSM 1:141; AT 6:59]). The surviving text never gets as far as describing the soul, not to speak of the union of the body and the soul. It only describes how a purely corporeal human-like machine could work. The metaphor of humanly constructed machines (clocks, artificial fountains, mills) is present right from the beginning of the text, and the given impression is that the human body could work without the soul just as well as these machines can.

At the end of the treatise, Descartes turns to the topic of the soul and indeed to the different kinds of soul. On the second to last page (of the CSM translation), he reasserts that he is now going to turn to the soul, but now he puts the point as "before going on to describe the rational soul" (CSM 1:107; AT 11:200). And indeed, as one can notice, he has already explained what would traditionally have been explained with reference to the vegetative or the sensitive souls. On the last page, he provides a list of the functions that have been explained. This list is worth quoting in full:

---

[2] René Descartes, *The Philosophical Writings of Descartes I–III*, trans. J. Cottingham, R. Stoothof & D. Murdoch (Cambridge: Cambridge University Press, 1984–1994), quoted as CSM 1–3; René Descartes, *Œuvres de Descartes I-XII*, rev. ed., ed. Ch. Adam & P. Tannery (Paris: Vrin/C.N.R.S., 1964–1974), quoted as AT 1–11.

> I should like you to consider, after this, all the functions I have ascribed to this machine – such as the digestion of food, the beating of the heart and arteries, the nourishment and growth of the limbs, respiration, waking and sleeping, the reception by the external sense organs of light, sounds, smells, tastes, heat and other such qualities, the imprinting of the ideas of these qualities in the organ of the 'common' sense and the imagination, the retention and stamping of these ideas in the memory, the internal movements of the appetites and passions, and finally the external movements of all the limbs (movements which are so appropriate not only to the actions of objects presented to the senses, but also to the passions and impressions found in the memory, that they imitate perfectly the movements of a real man). (CSM 1:108; AT 11:201–02.)

The list could be (perhaps one should say "probably is") based on an overview of scholastic textbooks discussing the functions of the vegetative and sensitive souls. Yet Descartes makes it clear that all these functions can be explained purely mechanically without reference to any soul. In his own words: "In order to explain these functions, then, it is not necessary to conceive of this machine as having any vegetative or sensitive soul" (CSM 1:108; AT 11:202).

What I find particularly important here is that Descartes thinks that he can also explain the functions of the animal soul for which the Latin word *cognitio* was traditionally used: external senses, common sense, imagination, memory, appetites, and passions. Descartes believes that he can explain how the cognitive process from sensation to action works mechanically when it proceeds purely on the level traditionally demarcated through the Aristotelian concept of the sensory soul.

With the explanation of these functions as mechanical, Descartes could be seen as separating physiology from psychology, to use twentieth-century terminology. This seems indeed to be the case to some extent. In the traditional Aristotelian model, the study of the soul would lump together all the functions of living entities *qua* living entities. Descartes argues instead for a system where the specifically human functions – thought and will – cannot become objects of study in the same way as the vegetative and sensorimotor functions. Because his dualism divides the very thing that was known as the soul into two radically distinct domains, no unified science of the soul in the traditional sense remains possible.

However, if we understand Descartes to be formulating the distinction between physiology and psychology in the way it was understood in the twentieth century, we fall into to a grave misunderstanding of Descartes' project. Many of the phenomena studied by Descartes as bodily events were in the twentieth century studied in psychology, not in physiology. For example, Jean Piaget's theory concerning the sensorimotor stage of human development is standardly classified in psychology because the child acts as a psychological entity already at this stage: his or her activity is related to his or her psyche.

Indeed, if it seems clear that there is no Cartesian "science of the soul" in the Aristotelian sense, it seems equally clear that there is no Cartesian "science of the soul" in the sense in which modern psychology is a science of the soul. The Cartesian soul is not like the Aristotelian soul, but neither is it like the modern psyche. Thus, if we accept that psychology studies all the phenomena to which the psyche pertains, we cannot refer these phenomena to the Cartesian soul.

This difference between the Cartesian soul and the modern psyche becomes apparent already in the passage from Descartes' *Treatise on Man* quoted above. There seems to be no denying that many of the phenomena explained by Descartes with reference to the mechanical body are psychological and would pertain to the psyche in the modern sense. That movements of the body are appropriated to external events through sensation, memory, and the arousing passions is a psychological fact, not a bodily or a mere physiological event in the modern sense. Even if passions would turn out to be reducible to physiological phenomena, they would still be studied at psychology departments. They are an essential part of the phenomena of the psyche, as the term is understood today, even in materialist theories of the mind. In general, modern psychology is open to and indeed often endorses the idea that the psyche it is studying might not transcend the body. But Descartes' point in the *Treatise on Man* is that the chain from sensation through passion to action can occur without the presence of a soul, that the soul would not be involved in this chain. His explanation involves concepts that today would have to be called psychological, but nothing in the explanation is about the soul.

Sight and other types of sensations are bodily phenomena in the Cartesian framework. Evidence to the contrary might be found in Descartes' listing of the different kinds of thought that we find in the *Meditations on First Philosophy* and some other connections. In the Second Meditation, he describes what a thinking thing is: "A thing that doubts, understands, affirms, denies, is willing, is unwilling, and also imagines and has sensory perceptions" (CSM 2:19; AT 7:28). Imagination and the senses are in the *Treatise on Man* listed as faculties that can be explained through the body only, without reference to the soul. So why are they modes of thought?

It is clear that Descartes did not change his mind on this issue between the *Treatise on Man* (written around 1632–1633) and the *Meditations on First Philosophy* (published in 1641). The letter to Regius from May 1641 discussed above is not the only proof of this. For, in the Second Meditation itself, about a page before the thinking thing is described, Descartes is in his doubts and considers nutrition, movement, and sense-perception in the following way, before moving on to thought:

> But what about the attributes I assigned to the soul? Nutrition and movement? Since I now do not have body, these are mere fabrications. Sense-perception? This surely does not occur without a body, and besides, when asleep I have appeared to perceive through the senses many things which I afterwards realized I did not perceive through the senses at all. (CSM 2:18; AT 7:27.)

Here, sense-perception is, along with nutrition, a thing that I may appear to have even when I in fact do not (note: "which I afterwards realized *I did not perceive through the senses at all*"; "animadverte *me non sensisse*"). It may be of interest to note that here Descartes uses the verb *sentio*, which does not involve the word *percipio* that is usually translated as "perceive" and is quite clearly connected to the mental sphere in Descartes' Latin. In this passage, no difference between nutrition, movement, and sense-perception is suggested. All are judged to be bodily operations that I may falsely dream to have, and in this respect they differ from thought – I cannot merely think that I think without actually thinking at that very instant.

And indeed, if we go back to the characterization of what a thinking thing is, we find further discussion on the status of sense-perception. In Cottingham's translation, the relevant lines are as follows:

> Lastly, it is the same 'I' who has sensory perceptions, or is aware of bodily things as it were (*tanquam*) through the senses. For example, I am now seeing light, hearing noise, feeling heat. But I am asleep, so all this is false. Yet I certainly *seem* to see (*videre videor*), to hear and to be warmed. This cannot be false; what is called 'having a sensory perception' is strictly just this (*hoc est proprie quod in me sentire appellatur*), and in this restricted sense of the term it is simply thinking. (CSM 2:19; AT 7:29.)

It seems to me that in this passage Descartes recognizes that the same doubting 'I' also has sensory perceptions. However, according to the methodic doubt, having these sensory perceptions is taken to be false. What remains certain is that the 'I' "seems to" have sensory perceptions, and this appearance is what is to be called sensory perception "in the 'I'" itself. Furthermore, this appearance of having sensory perceptions is a mode of thought and thus does not commit to having a body.

Now, does Descartes assign sensory perception to the mind here? I find it hard to read the passage in such a manner. The only indications of sensory operations that we can find in the 'I' which is involved in methodic doubt (and then lacks a body) are appearances of sensory perception. Cottingham skips in his translation the crucial pair of Latin words *in me* (roughly, "in me" in English), obviously for the sake of fluency – and because it is very hard to see what exactly their import is. Descartes is not, however, saying *tout court* that a certain appearance strictly speaking is called "having a sensory perception." Rather, he is saying that such appearance *in me* is called "having a sensory perception." As I see it, this could be paraphrased into claiming that insofar as we speak only about my mind, having a sensory perception involves no more than it seeming to me that I am having a sensory perception.

Normally, when we say that things seem to be a certain way, they can either be that way or not. As I read this passage, this is what Descartes means. When it seems to the meditator that he sees light, he either sees it, or he does not. In my view, the meditator's problem is not that he is unsure whether there is light or not, but, even more fundamentally, the meditator is unsure whether he has eyes or not. And if he does not have eyes, there is no way he can really see, even if it might appear to him that he sees.

The point is that, alone, the Cartesian soul does not have real sense-perceptions, because sensation is something that the body does. The soul does have awareness of the subject itself having sensations, and this may, of course, in many contexts – especially epistemological ones – be assimilated into sensation itself. In the extreme circumstance of radical doubt, the meditator is not, however, ready to accept that he is really having any sensory perceptions, and to this extent self-awareness may be illusory. Even if the meditator has awareness of himself seeing light, he might not be seeing at all. In essence, the structure is the same with many other bodily operations. For example, Descartes admits that walking is an acceptable basis for a cogito-like inference in the Fifth Set of Replies. The inference is "I am walking, therefore I exist". According to Descartes, this inference is acceptable "in so far as the awareness of walking is a thought" and even so only if "applied to this awareness,

and not to the movement of the body" (CSM 2:243–44; AT 7:351–52).[3] The comparison to sight is obvious: I could have awareness of seeing even if I did not really see, and similarly I could have awareness of walking even if I did not really walk. For thought itself, the situation is different: it could not happen that I have awareness of thinking without *ipso facto* thinking.

The crucial point to note here is that not only seeing but also walking is a bodily operation of which one is aware in a way that yields metaphysical certainty. It appears that the example comes from book IX of Aristotle's *Nicomachean Ethics* (1170a25–1170b1). Aristotle lists certain operations of the soul such that when one is performing them, one is aware of performing them and consequently also of one's own existence. It seems that Descartes admits that these operations are accompanied by a reflexive awareness and that existence (and, in the Sixth Meditation, also embodiment) can be inferred from this awareness.

Aristotle's list includes walking together with operations that we too would regard as immediately psychological: seeing, hearing, and thinking. It is important to bear in mind that there is no metaphysical separation between walking and seeing, neither for Aristotle nor for Descartes. Seeing is not a purely mental function any more than walking. Both are bodily functions that can be and typically are accompanied by self-awareness. Even if one of them is now more commonly regarded as psychological, there is no reason to suppose that this exact classification would have always been recognized in intellectual history. However, demarcating the field of psychology in terms of subject matter appears to require some consistent way of distinguishing psychological phenomena from other kinds of things. Nowadays, sight is taken to be a function of the psyche in a much stronger sense than walking: we are understood to see with our minds but walk with our feet. Neither Aristotle nor Descartes noted any distinction like that. Thus, they could not have recognized psychology as we know it.

4.2 Psychology in the Cartesian System of Sciences

At this point, let us turn to the Cartesian classification of sciences. If it seems difficult to find from Descartes a psychology like the one studied at the turn of the third millennium, we may have more success if we accept the Cartesian division of sciences and look at what would be the locus that comes closest to modern psychology.

The best overall presentation of the Cartesian system of sciences comes from Descartes' *Principles of Philosophy*, which was intended as a course-book giving an overall presentation of all the fields belonging to what was at that time called

---

[3] Descartes is very clear that walking is a function of the soul as much as seeing: "For shortly afterwards I expressly referred nutrition to the body alone; and as for movement and sensation, I refer them to the body for the most part, and attribute nothing belonging to them to the soul, apart from the element of thought alone" (CSM 2:243; AT 7:351). What exactly is the "element of thought"? In my view, it must be spelled out in terms of voluntary control and consciousness – and in these respects walking differs from nutrition in the same way as seeing.

philosophy. According to Descartes' own definition in the preface to the French version of *Principles of Philosophy*,

> [. . .] 'philosophy' means the study of wisdom, and by 'wisdom' is meant not only prudence in our everyday affairs but also a perfect knowledge of all things that mankind is capable of knowing, both for the conduct of life and for the preservation of health and the discovery of all manner of skills. (CSM 1:179; AT 9:2.)

In the seventeenth century, philosophy (*philosophia*) had a very wide meaning. It included everything that we would call "science", but also more. In particular, what we call "philosophy" was philosophy also at that time. In an interesting way, mechanics and ethics were also included in philosophy, and somehow Descartes seems to think of them together with medicine as a group on the same level.

Some pages later, Descartes presents the well-known analogue between the system of sciences and a tree. The idea is that a student who has gained sufficient skills through studying the art of thinking or logic and mathematics would then turn to "true philosophy in earnest". The tree, which the student would then climb, would have metaphysics as its roots, physics as the trunk and medicine, mechanics, and morals as the branches carrying the fruit (CSM 1:186; AT 9:14). In a more particular description of how the student ought to make progress, Descartes says that one should begin with metaphysics and study physics second. Turning to physics, or philosophy of nature, the student would then:

> [. . .] examine the general composition of the entire universe, and then, in particular, the nature of this earth and all the bodies which are commonly found upon it, such as air, water, fire, magnetic ore and the minerals. Next we need to examine individually the nature of plants, of animals and, above all, of man, so that we may be capable later on of discovering the other sciences which are beneficial to man. (CSM 1:186; AT 9:14.)

Of course, the system of sciences presented here is not very clear. The general setting, however, is visible, and interestingly enough it seems relatively easy to locate Descartes' own works in this system.

Here we need not dwell upon the methodological works (*Geometry*, *Rules for the Direction of Mind*). But when we start from the roots of the tree, it is important to recognize that Descartes' *Meditations on First Philosophy*, the first part of his *Principles of Philosophy*, and to a large extent also the *Discourse on Method* must be located as foundational, or at the roots of the tree. In the words of the preface of his *Principles of Philosophy*, the topics discussed are "the principles of knowledge, including the explanation of the principal attributes of God, the non-material nature of our souls and all the clear and distinct notions which are in us" (CSM 1:186; AT 9:14).

Now, is "the non-material nature of our souls" a topic of psychology? This depends on how we understand the expression. In the twentieth century, the relations between psychological processes and neuronal events in the brain were studied at psychology departments. In this sense, a part of psychology could perhaps be found within metaphysics, which Descartes describes as studying "principles of knowledge". (One might be tempted to say that this ought to rather be called epistemology.) Some recent commentators have even claimed that the Cartesian

study of clear and distinct notions is psychological in the sense that it consists of quasi-empirical studies of what we find ourselves forced to grant in our thoughts. The idea would then be that it is a psychological fact that you cannot deny that $2 + 2$ equals 4. If this is a psychological fact, then this area of psychology is studied in Cartesian metaphysics. Other scholars, however, think that Descartes does not confuse psychological and metaphysical compulsion but, on the contrary, keeps them carefully separate. According to them, insofar as clarity and distinctness are studied in Cartesian metaphysics, they are not studied as psychological topics in the twentieth-century sense.

In the above quoted description of the contents of physics, Descartes proceeds to the nature of living entities, plants, animals, and especially humans. This classification is reminiscent of the traditional Aristotelian study of the soul, which Descartes seems to adopt. In this system, the study of animate nature follows without disruption after the study of inanimate nature. Descartes is thus not unambiguously rejecting the traditional system of the sciences. But in order to properly understand how to locate psychological issues in this system, we must take a closer look at the contents of the Cartesian study of animate nature.

In the *Principles of Philosophy* part IV, § 188, Descartes apologizes for not including a fifth and a sixth part to his presentation. The fifth part would have discussed plants and animals and the sixth, humans. As the preface to the French version explains, his resources did not allow him to complete these sections within an acceptable timeframe. He simply did not have enough money for the necessary empirical work, although he did not "yet feel so old" that age would hinder him from doing the work (he was 51). Given certain other obstacles, he announces that he will continue philosophizing only for his own pleasure and leave the systematic completion of his system for future generations. So the systematic presentation we have in the *Principles of Philosophy* is disrupted at the point where Descartes would have turned to issues traditionally discussed in the Aristotelian science of the soul.

Nevertheless, the *Principles of Philosophy* contains a summary of some important issues that would have been discussed, and a number of unpublished texts discussing the relevant topics have survived. Furthermore, there seems to be good reason to suppose that Descartes' last published work, the *Passions of the Soul* is not unrelated. Proceeding in an orderly way, we must however first ask: What did Descartes write about plants and animals?

The answer is: not much. His published works include very little apart from a couple of pages in the *Discourse on Method*. There, Descartes spells out his famous comparison to man-made machines, claiming that irrational animals have no souls and that the rational use of language is the main attribute by which we can distinguish humans from other animals (CSM 1:139–41; AT 6:56–9). From the Sixth Set of Replies to the *Meditations on First Philosophy*, we can see that Descartes did not change his public stance (CSM 2:287–89; AT 7:425–27), and we find it defended also in his letters, most importantly in the letters to Henry More (CSM 3:366; AT 5:278).

Editions of Descartes' works contain a surviving manuscript entitled the *Primae cogitationes circa generationem animalium* ("First thoughts concerning the generation of animals"). Most probably the text is from Descartes' pen, but on many issues

it is so fragmentary that it is difficult to judge how strongly he would have defended the views discussed.[4] It seems even possible that some of the ideas have simply been copied from some other source for later personal use. But in any case, the text shows Descartes' interest in certain crucial issues concerning what life and especially animal life is like. There seems to be reason to suppose that this manuscript would have provided material for the part of science that would have been discussed in the fifth part of *Principles of Philosophy* if it were ever written.

For our purposes, the interesting parts of the text are those that shed light on the peculiar theory of animals as machines. Even in this text, Descartes is very clear about the metaphysical claim that animals have no souls. However, he makes it equally clear that the absence of the soul does not imply an omission of terminology that we would deem psychological. Animals have senses, and they can remember corporeal things; although they do not think – there is no mind. Most importantly, seeking advantage and avoiding harm guide their behavior. Sense-perception guides animals to enjoy things that are advantageous to them. In Descartes' own words:

> One ought not to wonder that the brain of brute animals there are so many different structures, since we see them move in so many ways. For all their movements arise from only two principles, advantage to the nature, or harm, and this to specific parts or to the whole. Thus, when senses present something advantageous to the whole, immediately this movement, which effects the senses, effects also in other members all the movements for enjoying these advantages. If they present something advantageous only to one part and disadvantageous to the others, this movement that is sensed, determines animal spirits to effect in one part all possible movements by which this advantage is enjoyed, and in others such by which this disadvantage is avoided. (AT 11:519; translation MY.)

In this passage, we can see certain important features of Descartes' machine-metaphor relatively clearly. Animals are like machines in the respect that their behavior is produced through a causal chain. As the text continues after this quote, Descartes explicitly brings up the fact that brute animals have no free will. Just like machines, bodies (and body parts) necessarily belong to causal chains (AT 11:519).[5] Indeed, they always act upon sensory perception or upon "movement given by nature". The lack of soul or mind implies that there is no power that could guide the animal in any other direction than that determined by the way in which sensory perceptions are processed in the brain. Animals, unlike humans, impeccably aim at enjoying advantage.

The machine-metaphor does not, therefore, imply that the concepts 'sensory perception' (*sensus*), 'advantage' (*commodum*), and 'enjoying' (*fruor*) would not apply to animals. On the contrary, Descartes specifically relies on such terminology in his explanations. This is one of the main issues emphasized in his letters to Henry More. In a letter to the Marquess of Newcastle, November 23, 1646, Descartes goes even further. He admits that animals have "some thought":

> The most that one can say is that though the animals do not perform any action which shows us that they think, still, since the organs of their bodies are not very distinct from ours, it may

---

[4] The manuscript is included in AT 11:501–38.

[5] See also *Principles of Philosophy* book I, § 37 (CSM 1:205; AT 8:18–9).

be conjectured that there is attached to these organs some thought such as we experience in ourselves, but of a very much less perfect kind. (CSM 3:304; AT 4:576.)

It seems relatively clear that for Descartes the machine-like character of animals (including humans with respect to their bodies) does not exclude attribution of psychological characteristics. On the contrary, it is exactly the corporeal, machine-like structure shared by humans with other animals that makes one think that animals would be like us in their thoughts as well – although they do not speak nor act freely.

In the manuscript on the generation of animals, Descartes discusses the structure of an oyster, pointing out the parts which correspond to the brain and certain other crucial organs (AT 11:520). After this example, he points out that brute animals do not have *knowledge* of advantage and disadvantage. The idea is that in order to seek its own advantage the oyster does not need to *know* what advantage is. Rather, the oyster needs to have since its organ of generation is in a place corresponding to the uterus some natural movement of its own, which is strengthened by certain types of movements and weakened by others. Unfortunately, this discussion about the mechanical character of how advantage is established in individual animals is very short. This is, however, perhaps the greatest philosophical mystery in Descartes' mechanical theory of the animal.

In the *Meditations on First Philosophy*, in the Sixth Meditation, Descartes puts forth a metaphor of a clock in such a way that it seems like he would deny all intrinsic teleology from all corporeal things. He considers the case of dropsy. In this condition, the body is sick. The sensory perception of a dry throat makes it seek drink which it does not need. Thus, the body seeks things which are disadvantageous to it. Descartes compares such a body to a badly-made clock, which does not tell the right time. Both the sick body and the badly-made clock deviate from their nature, but, as Descartes takes great pains to show, both follow all the laws of nature. In fact, there are two senses of the word "nature" at play, and Descartes seems to suggest that the nature from which the sick body and the badly-made clock deviate is mind-dependent.

How should we connect the passage in the *Meditations on First Philosophy* to the description of animal psychology in Descartes' unpublished works? The passage in the manuscript on the generation of animals clearly suggests that explanations of animal behavior must be teleological and involve psychological concepts, though not the concept of the mind. In the Sixth Meditation, Descartes seems to suggest that teleological terminology can be meaningfully applied only to the mind, or to the mind–body union. The implication would be that animal health or animal well-being is as mind-dependent an ascription as the clock's telling the right time.

There are at least two ways of connecting these passages. Either Descartes consistently acknowledges animal welfare, or he remains hesitantly self-contradictory about a topic to which he often returns in his unpublished works. I do not think there is a way of reading Descartes' letters and the manuscript on the generation of animals that would not acknowledge genuine animal welfare.

It does seem somewhat appealing to say that Descartes would accept that he is inconsistent here. This topic is central to his programmatic idea that all corporeal nature should be explained mechanically without reference to teleological principles. Then again, there seems to be little reason to ascribe a mind and rationality to an oyster, whose behavior appears teleological. Thus, a credible explanation of how animals fit into the metaphysical project seems to be contradictory, and, thus, it is no wonder that Descartes in his unpublished texts ponders about how to explain animal teleology and takes preliminary stances that are inconsistent with his published works.

To me, it seems more reasonable to read Descartes as acknowledging a significant difference between living and non-living things. True enough, he claims in a letter to Regius in June 1642 that this difference is no greater than the difference between self-moving machines and equipment like swords or keys that do not move by themselves (CSM 3:214; AT 3:566). A suitable point of comparison seems to be coming forward from certain forms of twentieth century enthusiasm concerning artificial intelligence. It may have seemed to Descartes that an important, almost metaphysical, barrier is broken when humans have managed to produce machines that move themselves (almost) perpetually and are not dependent on an external mover. With some further engineering, perhaps a self-adjusting clock could be made so that it would heal itself from occasional malfunctions and return to showing the right time. This could, perhaps, be called genuine teleology despite its not being intrinsic. Also, it might be that this would have appeared to Descartes like producing a living entity. After all, showing the right time is something that does not depend on human intentions despite the fact that the clock itself is not able to *know* the right time or even what it is to show the right time. The earth rotates at its own pace and determines the right time quite mind-independently, and a clock is well-made only if it catches this right time. To understand the relevant concepts, a mind is of course needed. In this sense it is clear that being a healthy body or being a well-made clock are said "from the outside," or technically speaking, they are extrinsic denominations (*denominatio insela*).[6] This does not, however, make the teleology any less real.

4.3 Cartesian Science of the Human Being

As we have now seen, the Cartesian theory of animals as machines would have included a number of issues that would have to be called "psychological" in a twentieth century context. It seems, thus, that a Cartesian psychology would be a science that concerns the corporeal world, at least to a large extent. Consideration of the *Treatise on Man* and certain other, later, texts strengthens this impression. Among these texts, is the *Description of the Human Body*, which most scholars think Descartes refers to in a letter to Elisabeth on January 31, 1648, stating that

---

[6] See the Sixth Meditation (CSM 2:59; AT 7:85).

he has been writing a work on "the functions of animals and man" (CSM 3:329; AT 5:112). In this text, Descartes' philosophical claims are more modest than in the early *Treatise on Man*, but it is easy to see that the program has not changed. Functions traditionally ascribed to the animal soul should be explained without reference to the soul even in humans.

The *Passions of the Soul* is from a psychological perspective clearly the most interesting text written by Descartes. Its topic is obviously psychological, and it contains a number of passages in which Descartes shows genuine psychological insight. In this sense, this work ought to reflect how a Cartesian psychology would look. Here we cannot engage in a full study of the work. Instead, let us satisfy ourselves with some general remarks concerning the general approach of the work.

Descartes starts with a careful distinction between the soul and the body, and he aims at giving an elaborate definition of passion (or "emotion", as we might translate the word). The definition makes it clear that in order to understand human emotions, we need to take into account the interactions between the soul and the body and, indeed, consider the human being as a unity. A large part of the work is dedicated to seemingly physiological explanations of bodily phenomena constituting the various kinds of emotions. Towards the end, the reader is expected to come to a deeper understanding of all the particular passions. As the two last articles of the text tell, the objects that cause the passions produce the bodily effects so rapidly that human wisdom would not be able to eradicate all the passions. Indeed, what one should rather do is to learn to live with the passions. For the positive ones, this consists mainly in the ability to control one's actions even when mislead by excessive passion. However, the case of negative passions is of course more difficult. But, as the last sentence of the work claims, with wisdom the evils caused by passion can be bearable "and even become a source of joy" (CSM 1:404; AT 11:488).

As mentioned above, in the *Principles of Philosophy* Descartes compares the system of sciences to a tree. If we look at the *Passions of the Soul* in terms of this tree, it seems that the book aims at showing the student the way from the trunk of the tree to certain branches and to the fruits. The book starts with physics and aims at moral conclusions concerning how one should live. The basics are of course very sketchy and the main emphasis is put on a discussion of how the human body functions in producing emotions in the mind. Given the Cartesian stance that animals have emotions but no minds, it seems that in the genuinely psychological part of the presentation, the soul plays only a limited though important role. Within the Cartesian framework, moral conclusions require attention to the soul. Only the rational human soul is free and thus capable of making moral choices. And it is through this mental mastery of the passions that the link between morality and the passions is found.

One of the crucial features of the passions discussed often in the work is their ability to prompt the soul to make some specific choice. Simply put, if you are afraid, you tend to flee. The actual choice of fleeing is, according to Descartes, under voluntary control, but it is equally clear that the fear caused by your bodily systems has an effect on your choice. For example, in part I, § 47, Descartes discusses internal conflicts as occurring between the soul and the body (CSM 1:345–46; AT

11:364–65). While the passions (the body) push your mind in one direction, your rational choice would take you in the other direction. As Descartes recognizes, it is unrealistic to claim that the rational mind is always stronger in these situations. Rather, one has to learn psychological techniques that operate on the bodily level. For example, through imagining a prize gained by being brave one can mollify fear and thus become able to act against it. As Descartes explains, this technique works, because the image of the prize in the brain produces movements in the animal spirits that counter the movements which produce the fear.

In his work on the passions, Descartes assumes that much of the cognitive processing responsible for one's behavior happens unconsciously, outside the soul. He uses the traditional Aristotelian idea that in normal life most of what we do is determined by systems that are common to humans and other animals. Sensory perceptions produce emotional responses in the brain and lead us to instinctual or customary behavior. No mind is needed for these psychological processes. However, a mind is needed for voluntary conscious control of them – for transcending our animal nature into the moral sphere. Thus, the mind belongs to the field of ethics rather than psychology.

4.4 Conclusion

The core part of psychology would in the Cartesian framework proceed as a science of the body. As a science of purposeful behavior, psychology studies the body, and it can treat humans and other animals practically on a par, at least insofar as moral motivation for action is not taken into account. As the study of consciousness and of mental phenomena, psychology would not study the body. It seems, thus, that there could not be a unified Cartesian psychology.

In the *Meditations*, in the Fourth Set of Replies, Descartes gives a list of examples of our movements which "do not depend in any way on the mind." This list includes: "heartbeat, digestion, nutrition, respiration when we are asleep, and also such waking actions as walking, singing and the like, when these occur without the mind attending to them" (CSM 2:161; AT 7:229–30). The example of singing is interesting. It seems that here Descartes has in mind the traditional idea that you may sing without paying any attention to the words or even understanding them.

Think about a person singing in the shower. She may be mentally concentrated on choosing the right shampoo but singing without even noticing that she sings, simply because the melody fits the sound of the falling water. She never consciously began to sing, the song just started to flow with the water. This is the kind of situation that Descartes seems to mean. The mind is elsewhere, but the singing continues. For the purposes of this chapter, it is important to note that if we want psychological explanations of the singing, or of the psychological processes involved, there seems to be every reason to suppose that from the Cartesian perspective these explanations must concern the body (mainly the brain), not the mind. In contrast, the mind plays the primary role in the explanation of how the right shampoo is chosen if the

person attends to her choice. The two operations, singing and choosing the shampoo, receive radically different explanations.

The example shows how the Cartesian theory distinguishes between different explanations of behavior very differently from twentieth century psychology. Descartes sees a distinction between higher and lower functions where we would not, at least not easily. Descartes' divide remains very close to the traditional Aristotelian line between the rational and the animal soul. This divide is something that is very difficult to see from our viewpoint. Then again, our division between psychological and physiological functions would probably seem strange to Descartes.

How should we, then, think about these distinctions? There is no obvious metaphysical reason to prefer the distinction between physiological and psychological functions over the Cartesian (or Aristotelian) division between rational and sensitive functions. Both should perhaps be considered as historical contingencies and, thus, as subject of examination. The unity of psychology as it stands today seems as accidental as the lack of any such unified discipline within the Cartesian framework.

Chapter 5
Imagination and Reason in Spinoza

Theo Verbeek

Abstract The chapter examines Descartes' successor and critic, Baruch Spinoza. Emphasis is put on clarifying the relation between three concepts that are central to Spinoza's account of our knowledge and understanding of the human mind: the concepts of intellect, imagination, and reason. For Spinoza, the highest form of cognition is the intuitive understanding, which the mind is able to achieve with its intellect. This form of knowledge concerns particulars, and it comprehends their adequate essences. Whereas imagination is a passive faculty that registers what goes on in the body, the ideas of the intellect result from the activity of the intellect itself. In Spinoza's account, reasoning and reason-based knowledge remain second in relation to intuitive knowledge. Reason provides general and abstract ideas which do not indicate any real existing entities but merely take part in the explanation of their relations, interconnections and differences, agreements and disagreements. By investigating these crucial distinctions, it is demonstrated how Spinoza struggles to change the ontological status Descartes gives to general notions, such as 'substance' and 'attribute' which "extend to all classes of things." According to Spinoza, general notions are neither existing objects nor properties of such objects, but they are *entia rationis*. In the final analysis, general notions turn out to be *entia imaginationis*, products of imagination, and it is concluded that Spinoza's notion of reason is in fact closer to the imagination than to the intellect.

Keywords General notions · imagination · intellect · reason · Spinoza · Descartes

According to Spinoza's letter to Johannes Bouwmeester (1666), "method" consists in the knowledge of the nature and laws of the intellect. All one has to do to acquire that knowledge would be to "make a distinction between the intellect and the imagination, that is, between true ideas and all others, namely, fictitious, false,

T. Verbeek
Department of Philosophy, Utrecht University

and doubtful ideas."[1] Similarly, according to the *Ethics*, if we are to understand the nature of error two distinctions are indispensable: (1) the distinction between ideas and images; and (2) the distinction between ideas and words – which, given the fact that words and language are part of the imagination, amounts to the same distinction.[2] So apparently the distinction between the imagination and the intellect is among the most important of Spinoza's philosophy.

Nonetheless, Spinoza hardly provides any definition either of the imagination or of the intellect. According to the *Treatise on the Improvement of the Intellect*, the imagination may be defined in any way we like as long as it is "that by which the mind is passive."[3] Its only role is to register what is going on in the body – the laws of association, by which imagination and memory are governed, are ultimately physiological laws.[4] The intellect, by contrast, is defined only by its opposition to the imagination, so it should be essentially active, which is indeed confirmed by a grammar dominated by the word "to form" (*formare*). So the ideas of the intellect are the result of an activity of the mind. The intellect is the mind in so far as it forms ideas. If the imagination forms ideas at all, it is against our will (*nobis invitis*).[5] By itself and without any constraint on behalf of the intellect, the imagination produces any number of complex ideas – as can be seen in prophets and presumably also in dreamers and lunatics.[6]

In my book on Spinoza's *Theologico-Political Treatise*, I deal with the social, political, and religious significance of this theory.[7] My subsequent work on the *Treatise on the Improvement of the Intellect* not only confirms the importance of Spinoza's theory but also leads to a clearer view of its relation to Cartesianism.[8] But many problems remain, especially with respect to the notion of 'reason', which according to Spinoza we should distinguish from the intellect. Generally speaking, the intellect is a source of intellectual intuitions, so its object would be particular, whereas reason is the faculty of forming and manipulating general and abstract ideas.[9] The aim of

---

[1] Spinoza to Bouwmeester, 10 June 1666, Ep. 37, IV 189. Apart from the *Short Treatise* (see below) Spinoza is quoted in the standard edition of Gebhardt (*Opera*, 5 vols., Heidelberg: Winter, 1925–1938). *Ethica* (Ethics) is cited as Eth; *Tractatus de intellectus emendatione* (Treatise on the improvement of the intellect) as TIE; *Korte Verhandeling* (Short Treatise) as KV; *Tractatus theologico-politicus* (Theologico-Political Treatise) as TTP; *Epistolae* (Correspondence) as Ep; and *Principia Philosophiae Cartesii more geometrico tractata* (The Principles of Philosophy by Descartes presented in a geometrical way) as PPC. All translations are my own.

[2] Eth II, prop. 49, schol., II 131.

[3] TIE §84, II 32.

[4] Eth II, prop. 18, schol., II 106–07.

[5] TIE, §108, II 38–9.

[6] For prophets see TTP i, III 28.

[7] Theo Verbeek, *Spinoza's Theologico-Political Treatise: Exploring "the Will of God"* (Aldershot: Ashgate, 2003), quoted as *Will of God*.

[8] Spinoza, *Verhandeling over de verbetering van het verstand*, transl. Theo Verbeek (Groningen: Historische Uitgeverij, 2002). The relation to Descartes and Cartesianism is also discussed in Ch. 6 of *Will of God* and in the article cited in note 13.

[9] Needless to say that Spinoza denies the existence of faculties; Eth II, prop. 48, schol., II 129.

this chapter is to clarify the nature of reason to some extent. But first we must deal with the imagination.

5.1 Imagination

According to Spinoza, the mind "imagines" whenever it represents a body as actually present.[10] However, if we perceive or imagine an external body what we actually perceive is our brain as it is (or as if it is) affected by an external body.[11] And we cease to believe in the presence of that external body only because the brain and nervous system are affected in a way we feel is incompatible with the existence of that external body.[12] So if the external body is no longer present, but the brain is not affected in a different way, we continue to believe in the actual presence of that body, that is, we "imagine" it in the more precise sense of the word. If I have a sensation of red, I believe that there actually is something red until the sense of touch reveals that such is not the case; if I dream, I believe in the reality of my dream as long as I do not wake up; if for a moment I think I see a ghost or a UFO, I discount that perception as a delusion, because I firmly believe there can be neither ghosts nor UFOs; if I believe that there is a God-Lawgiver, I shake off that belief as soon as I understand the full implications of the fact that "God" is infinitely perfect, etc. Inversely, I continue to believe I see something red if I do not verify my perception by the sense of touch; I continue to believe I see a UFO if I never ask questions about UFOs, etc.

As a result, the problem concerning the truth of the imagination is uncommonly complex. An idea of the imagination is true, that is, the existential judgment that accompanies that idea is true, if the imagined object really exists; it is false if the imagined object does not exist. But how do we know an object exists apart from our having an impression of that object? In itself a true idea of the imagination is not recognizably true or false – to assess its truth or falsehood we need either 'signs', that is, evidence which confirms or falsifies the implied belief in the existence of the imagined object, or 'norms', which are ideas that are true as well as known to be true (which cannot be ideas of the imagination).[13] So there is no *intrinsic* difference between a false and a true idea of the imagination; in itself a false idea of the imagination looks exactly like a true idea of the imagination and will be held to be true *cæteris paribus*. In any case, even a true idea of the imagination can never be known to be true absolutely. The fact that it is compatible with an idea that is true and known to be true means at best that it is *possibly* true and *morally* certain.[14]

---

[10] Eth II, prop. 17, schol., II 106; prop. 26, cor., dem., II 112.

[11] Eth II, prop. 16, cor. 2, II 104.

[12] Eth II, prop. 17, II 104.

[13] See Theo Verbeek, "'Verum index sui et falsi': shinri-no Kihan to Kijun" (in English 'Verum index sui et falsi': Norms and Criteria of Truth), *Spinozana: Supinoza Kyokai Nempo*, 4 (2003), 5–29 (so far available in Japanese only). For "signs" see *Will of God*, Ch. 3, 71–4.

[14] On the notion of 'moral certainty' see *Will of God*, Ch. 3, 75–81.

The only real certainty we can eventually achieve with respect to a given idea of the imagination is that it is certainly false – it is certainly false if it is incompatible with something that is true and known to be true, that is, to an idea of the intellect (like that there can be no God-Lawgiver).[15] This sets the record straight as far as the relation between the imagination and the intellect is concerned. But what about reason, that is, the faculty of discursive reasoning?

5.2 Reason

The picture presented of reason in the *Treatise on the Improvement of the Intellect* (where reason figures as "perceptions of the third kind") is more or less straight-forward.[16] With the help of general principles and laws, reason produces ideas of hypothetical causes, which, being general concepts, are 'inadequate' – indeed, apart from the fact that an idea can never be more true and more certain than the princi-ples and laws from which it is deduced, a general concept always relates to a class of beings, so leaves out certain things that may be characteristic for the thing that actually is the cause of a given phenomenon. By contrast, "perceptions of the fourth kind" (corresponding to what the *Ethics* calls "intuitive science") "comprehend the adequate essence of a thing," so are themselves adequate.[17] This is compatible with the *Short Treatise*. According to that work, reason (*rede*) shows how things must or should be, so it does not provide an immediate and intuitive awareness. And this also has practical consequences, because it means that reason cannot be the basis for moral action.[18]

---

[15] *Will of God*, Ch. 1, 24–6; Ch. 6, 158–62.

[16] Admittedly, Spinoza's description of perceptions of the third kind is vague and obscure, espe-cially if combined with his examples (which are confusing to say the least). What I deduce from it is, (1) that they are general and abstract; (2) that they explain given phenomena by means of a hypothetical cause; (3) that they may be arrived at in a way that is not absolutely safe (precisely because of their abstract nature); 4) that they are useless for salvation; see TIE §§ 19–28, II 10–3.

[17] TIE §19, II 10; §29, II 13. What Spinoza means, I suppose, is that we have a perception of the fourth kind whenever we understand something as an "adequate cause", that is, as "a cause such that its effects can be perceived clearly and distinctly through itself" (Eth III, def. 1, II 139). So if we have a perception of the third kind whenever on the basis of an observed phenomenon we deduce (with the help of a general law or principle) the properties the thing that caused it certainly had (so see the thing primarily as a member of a class), a perception of the fourth kind allows us to predict which effects a given thing must produce. This lack of symmetry between the explanation of an effect and the prediction of a cause is lifted by Hume's analysis of causality. Only a limited number of items can be adequate causes: God (TIE §38, II 15–6; §76, II 29) and our mind in so far as it is the cause of its ideas, that is, in so far as it is intellect (TIE §105, II 37; cf. Eth II, def. 3, II 85; III, def. 3, explic., II 139; V, prop. 31, II 299).

[18] The fact that reason does not provide immediate and intuitive awareness means that an idea of reason is always weaker (in the sense of "less practically effective") than an idea of the imagination (unless of course that idea of reason is reinforced by a passion – desire, for example – or by intellectual ideas). To have an idea of "the good" therefore is not enough to produce good actions. KV II (1–4), 42–51; (18–9), 82–92. The *Short Treatise* is quoted in the edition of Filippo Mignini (L'Aquila: Japadre, 1986).

In the *Ethics*, the picture is less straightforward. According to Spinoza, the mind reasons scientifically, not if it allows itself to be led "by the fortuitous course of events" to study this or that thing on its own, but if by studying several things together it tries to understand their agreements, differences, and incompatibilities.[19] Accordingly, Spinoza's view of reason (and science) has a Baconian ring: we collect data that are ordered and classified so as to constitute a "history", which in turn allows a systematic comparison of things, the result of which are "laws" and "forms". In a similar way, science is described in the *Treatise on the Improvement of the Intellect*: We must know nature in general as much as is necessary "to assess the differences, agreements and incompatibilities of things in order to form a true notion of what they are capable of and what not, and we must compare that result with the nature and capacity of man."[20] And also according to the *Theologico-Political Treatise*, science starts with a "history", from which in a second stage – called as by Bacon "interpretation" – definitions are drawn.[21]

According to the second *scholium* of prop. 40 of Pt II of the *Ethics*, however, "reason or knowledge of the second kind" is based on "common notions and adequate ideas of properties of things" (*notiones communes rerumque proprietatum ideas adaequatas*).[22] So science would (1) be based on "common notions"; (2) study properties of things rather than things themselves; and (3) work with "adequate" ideas. This seems to make a lot of sense. Usually "common notions" is just a synonym for "axioms" – it is in this sense at any rate that the phrase is always used by Descartes and often by Spinoza.[23] So science works with axioms, that is, presumably, with principles like "nature operates in a uniform way," or "every spatio-temporal event has a cause that can also be located in space and time," etc. That science does not deal with individual things is not controversial either. The only problematic claim is that the ideas of reason are "adequate", whereas in the *Treatise on the Improvement of the Intellect* Spinoza specifically claims that the ideas of reason are not *adequate*. This problem can be solved by assuming that, whereas in the *Ethics* Spinoza is interested in properties, in the *Treatise* he is interested in ideas of things – so in the *Treatise* his claim would be that an adequate idea of a thing's properties does not entail the adequate idea of that thing or its essence (which is not implausible).

More severe problems arise if we look more closely. Allegedly, the object of the complex argument stretching from propositions 38 to 40 is the concept of 'common

---

[19] Eth II, prop. 29, schol., II, 114; cf. TIE §84, II 32; §91, II 34.

[20] TIE §25, II 12.

[21] TTP vii, III 98; cf. TIE §§99–102, II 36–37. See *Will of God*, Ch. 4, 99–105.

[22] Eth II, prop. 40, schol. 2, II 122.

[23] "C'est bien une notion commune de penser que, si une nature intelligente est indépendante, elle est Dieu," see letter from Descartes to Mersenne, 15 November 1638, in René Descartes, *Oeuvres de Descartes*, ed. Ch. Adam & P. Tannery, new edn. (Paris: Vrin/C.N.R.S., 1964–76), vol. 2, 435, hereafter cited as AT followed by volume and page numbers; "il est certain qu'on ne doit recevoir pour notion que ce qui ne peut être nié de personne" (Descartes to Mersenne, 25 December 1639, AT 2:629); "il me semble que c'est une notion commune très évidente que, *quod potest plus, potest etiam minus*, aussi bien que *totum est majus sua parte*" (Descartes to Mesland, 2 May 1644, AT 4:111); cf. *Meditations*, Second Set of Responses (AT 7:164–65); *Epistola ad Dinetum* (AT 7:580); Spinoza, PPC, praef., I, 127; Eth I, prop. 8, schol. 2, II, 50; TTP vi, III 88; vii, III 99; xiv, III 179.

notions'[24] – indeed, the *scholium* that closes this group opens with the words: "In this way I have explained the ground (*causam*) of those notions which are called *common* and are the foundations of our reasoning (*ratiocinii nostri*)."[25] But the rest of the *scholium* suggests that the concept of 'common notions' is only one of Spinoza's concerns. Apparently, all general ideas are formed in a regular and uniform way, even if everybody forms them in a different manner – indeed, all essentially belong to the imagination which, in so far as it is primarily based on perceptions of individual bodies and associations of ideas, is basically idiosyncratic. So presumably the general purpose of propositions 38–39 is to provide a blanket explanation, not only of common notions, but of general and abstract ideas (and/or principles) generally and to lay out the conditions of their use.[26]

5.3 Propositions 38–39

Let us first concentrate on proposition 38. In this proposition, Spinoza claims that those properties or aspects (he uses the unspecific *illa*) "which are common to all and are equally in a part as well as in the whole" cannot be conceived but adequately.[27] According to the corollary, which is not proven separately, "there are certain ideas, that is, notions, common to all men; for all bodies (lemma 2[28]) are similar in certain ways, which consequently (because of the preceding proposition) must be perceived by all adequately, that is, clearly and distinctly." So the fact that certain properties of bodies involve the concept of extension (lemma 2) makes it possible to narrow down proposition 38 to one concerning common notions. This is already fairly obscure but the demonstration of the proposition seems to be a typical example of *obscurum per obscurius*:

> Suppose *A* is common to all bodies and is equally in a part of each as well as in the whole, then I claim that *A* cannot be conceived but adequately. For its idea will be necessarily adequate in God, not only in so far as he has an idea of the human body but also in so far as he has ideas of each of those affections which involve in part the nature of, not only the human body, but also external bodies. So this idea will be necessarily adequate in God in so far as he constitutes the human mind, that is, in so far as he has ideas that are in the human mind. As a result, the human mind necessarily perceives *A* adequately and, not only

---

[24] The literature on 'common notions' in Spinoza is not very satisfactory, if only because they take "common notions" in a sense which does not seem to be Spinoza's own. See H.A. Wolfson, *The Philosophy of Spinoza*, 2 vols. (Cambridge Mass.: Harvard University Press, 1934), II, 117–50 (despite a few interesting points); Martial Gueroult, *Spinoza* (Hildesheim: Olms, 1974), II, 324–52.

[25] Eth II, prop. 40, schol. 1, II, 120.

[26] The meaning of prop. 40 is obviously to stipulate that if we use them in the right way, the things we deduce from them are as "adequately" understood as themselves.

[27] Eth II, prop. 38, II 118.

[28] The text of this *lemma* is: "All bodies agree in certain respects" (*omnia corpora in quibusdam conveniunt*). The demonstration points out that this is the case, because they all involve the concept of one and the same attribute.

in so far as [the mind] perceives itself, but also in so far as it perceives either its own or an external body. Nor can *A* be conceived in any other way.[29]

The description of *A* as something that "is equally in a part of each as well as in the whole" is reminiscent of the traditional description of a *universal* as something that can be predicated of one as well as of many. But here we are not dealing with classes (*multa*) but with wholes (*tota*), that is, if we adopt the definition of the *Short Treatise*, things or beings "whose parts are non-discrete indivisible entities, which may belong either to the same or to different kinds" – whereas a universal is the concept of a class consisting of discrete indivisible entities which belong to the same kind.[30] So a *totum* is a *continuum* whereas a class consists of separate things.

This also seems to be the difference in the next proposition, which deals with wholes as well as classes:

> That which is common and proper to the human body and certain external bodies by which the human body uses to be affected and which is equally in a part of each as well as in the whole, of that, too, there will be an adequate idea in the mind.[31]

Whereas in the earlier proposition we were dealing with a whole and its parts, we are now dealing with discrete wholes (one of which is the human body), having certain properties or aspects in common, which in turn are proper, not only to them but also to their respective parts. Moreover, given the reference to lemma 2, the first proposition deals, not with any particular whole, but with the material universe (which, since there is no void, forms a continuum), whereas the second proposition deals with the macroscopic world as it is normally perceived by a human mind, namely, as being composed of discrete beings which may, or may not, have certain things in common, whilst at the same time taking those beings as wholes (and not as composites).

According to Spinoza, the properties dealt with in proposition 38 (*A*) are properties equally distributed through a whole and common to all *bodies*. So *A* is what we would call a fundamental property of matter, that is, a property found in any material body. The idea of *A*, moreover, would be necessarily adequate. Why? The only reason I can think of is that *A* is a fundamental property of matter and that therefore the idea of *A* is a true idea. This also seems to be the point of the demonstration: Given the fact that *A* is a quality common to all bodies, God, in knowing the material universe, has the idea of *A*; so to the extent that we also have the idea of *A*, our idea of the world is the same as God's and for that reason is adequate.[32] But on close reading, it is just the reverse: The idea of *A* is adequate in God, not because all God's ideas are adequate, but "in so far as [God] constitutes the human mind, that is, in so far as he has ideas that are in the human mind." And this entails, apparently,

---

[29] Eth II, prop. 38, dem., II 118–19.

[30] KV, second dialogue, ed. Mignini, 21. Both wholes as well as universals are said to be *wezens van reeden*, that is, *entia rationis*. Spinoza here seems to think of "universals" in the sense of class concepts ('man,' 'animal,' etc.) rather than universal properties.

[31] Eth II, prop. 39, II 119.

[32] This is also the way this is interpreted by, for example, Martial Gueroult.

not only a perception of *A* as it occurs in the material universe (that is, a sensation of *A*), but also a perception of the mind by itself: "The human mind necessarily perceives A adequately and, not only in so far as [the mind] perceives itself, but also in so far as it perceives either its own or an external body." Accordingly, the idea of *A* is adequate if and only if it entails a perception of the mind by itself, for it is only to that extent that God has the idea of the human mind thinking *A* and not simply of *A*. This would mean that God would not have the idea of *A* if there were no human mind having an idea of *A* – God perceives *A* only in so far as there is a human mind perceiving *A*. Does that make sense?

Spinoza usually calls ideas that involve a perception of the mind by itself "ideas of ideas".[33] Most commentators take this to mean "reflexive consciousness".[34] Accordingly, if I have a sensation of red (an idea of the first degree), I would have the idea of having the idea of red (an idea of the second degree) whenever I am conscious of having a sensation of red; and whenever I am conscious of the fact that I am consciously having a sensation of red I would have the idea of the idea of the idea of red (an idea of the third degree), etc. But this telescopic multiplication of ideas can be epistemologically relevant only if we interpret "idea" as "concept" rather than "awareness".[35] This means that if I have the idea of an idea, I am not simply conscious of having an idea but form a concept of that idea. So, for example, if I have a sensation of red, I not only have an idea of 'red' but, as we have seen, I also believe that there is a red object in front of me – and not only do I consciously believe that there is that object (say, a red flower); I also believe that that object is really red (I do not detach 'red' from the flower to which it belongs). All this would be an "idea of the first degree," provided we are allowed to call all of this an "idea". In a next stage I may realize that I can have a sensation of 'red' even if there is no red object around (that I can remember, imagine, dream or hallucinate a red flower); that I can detach the quality of red from any kind of material object to study it in itself – so the idea of the second degree would be the universal concept 'red' or, even more universal, 'colour'. And finally I can propose a theory of light, according to which colours exist only in the mind. The last two stages (which could be further analyzed) can certainly be described in terms of ideas of ideas, that is, in terms of a growing realization that what is given to me is not objectively there but is "something in my mind" – ideas of ideas necessarily involve an idea of 'mind'. As a result, science would presuppose ideas of ideas, not in the sense that I am aware of what I am seeing and hearing, but in

---

[33] Eth II, prop. 29, II 113; cf. TIE §33–34, II 14.

[34] This is also suggested by Eth II, prop. 43: "Someone who has a true idea knows that he has a true idea and cannot doubt its truth" (II, 123). But this refers to true ideas, that is, to ideas known to be true (which, of course, by definition involve the idea of an idea). A different interpretation is adopted by Bernard Rousset in his commentary of TIE, see Baruch Spinoza, *Traité de la réforme de l'entendement*, ed. B. Rousset (Paris: Vrin, 2002), 221–33, but I am not sure that I understand what he means.

[35] Eth II, def. 4, II 84.

the sense that I cannot study a quality before isolating that quality as an idea in my mind.

Now an idea of an idea is as real and as intelligible a 'thing' as any other thing of nature.[36] Accordingly, an idea of an idea is as necessarily 'willed' and 'conceived' by God as any other thing or event. However, unlike a sensation, which reflects only a particular state of a body and is therefore part of whatever God 'wills' and 'conceives' with respect to the material order of things, ideas of ideas are not only produced by my mind, but also obey a different type of laws (the laws of logic rather than the laws of association).[37] Accordingly, it is only to the extent that I conceptualize my sensations, that is, to the extent that I identify something as belonging to my mind, that God can be said to have, not simply an idea of my body (which idea coincides with my soul[38]), but an idea of my mind, for it is to this extent only (and not to the extent that my mind just registers the state of my body) that my mind *is* something. So, if for an idea to be adequate, it is necessary to be the idea of an idea, the idea of *A*, that is, the idea of a fundamental quality of matter, then it is adequate only to the extent that I realize that it is an idea, not to the extent that I am aware of the fact that this or that body is extended or is in motion. Accordingly, for my idea of *A* to be adequate it is not enough for me to perceive this quality but to form an idea of it, for it is only "as an idea" that this quality can become the object of a particular science (mathematics, mechanics, etc). This presumably is also the explanation of the corollary on 'common notions'. To the extent that we study qualities like extension and motion "as ideas" (that is, not as properties which belong to *particular* bodies), that is, to the extent that we are doing mathematics or mechanics, we also need the axioms of mathematics and mechanics.

In any case, the reason why the idea of *A* is adequate is not that it is the way God knows the world. The idea of *A* is not part of God's knowledge of the world as such, especially if *A* is a universal. Universals do not exist, so are not perceived by God either, except to the extent that he perceives *us*, that is, our minds, having the idea of a universal.

If we now turn to proposition 39 it becomes clear that what is meant here cannot be the fundamental qualities of matter meant in proposition 38. Could it be that in proposition 39 Spinoza means secondary qualities like 'red' or 'hot'? To the extent that in proposition 39 Spinoza deals with the macroscopic world, this could be the case, but it would create another difficulty. For not only are secondary qualities not real qualities; the secondary qualities of a whole are not necessarily those of the parts that make up that whole – the fact that my body has a certain colour does not entail that all parts of my body have the same colour; the fact that a body has

---

[36] TIE §33, II 14.

[37] According to Spinoza, laws of association are laws of the body which reflect the structure of the brain; Eth II, prop. 17–18, II 104–07.

[38] Eth II, prop. 13, II 96. I am not sure, for that matter, that this is the correct interpretation of that text; indeed, I am inclined to believe it means that the soul "exists" only to the extent that there is an awareness of the body.

a certain weight does not entail that its parts have the same weight, etc.[39] But are qualities meant at all? An alternative and more interesting candidate is suggested in *Metaphysical thoughts* (1663).

5.4 Beings of Reason

Spinoza's general point in part I of this work – devoted to general metaphysics or what we would call "ontology" – is that we must make an unambiguous distinction between being and not-being which, moreover, exactly coincides with that between necessity and impossibility. Accordingly, something either is or is not. If it is, it is necessary; if it is not, it is impossible. If it is, it can be conceived; if it is not, it cannot be conceived at all.[40] It is against this background that Spinoza discusses the notions of *chimaera* (something that cannot possibly exist because it unites two or more *manifestly* contradictory terms), *ens fictum* (an idea deliberately – *sciens et prudens* – contrived, which may, or may not, contain a *hidden* contradiction) and *ens rationis* (something devised by the mind in order to retain, explain, and imagine the things it has understood). For our purposes, we can discount the *chimaera*, which is not even an idea but a purely verbal being.[41] We can also discount the *ens fictum*, for on closer study an *ens fictum* will turn out to be either a real being (*ens reale*), or a *chimaera*, that is, a manifestly false idea.[42] In fact, *entia rationis* are by far the most interesting in this context.

According to Spinoza, *entia rationis* help us to remember, explain, or imagine the things we understand. They are explicitly said not to be *ideas*, having "no *ideatum* that exists or could exist." As a result, an *ens rationis* is a *modus cogitandi* only, that is, it is internal to the human mind.[43] Accordingly, *entia rationis* are known to God only in so far as he preserves and knows the human mind.[44] The fact that *entia*

---

[39] Apart from this, if secondary qualities are meant, it is impossible to make sense of the corollary attached to proposition 39, which claims that "the mind is more apt to perceive more things adequately according as its body has more in common with other bodies" (II 119).

[40] CM I, 3, I 240–44. Of course these claims can be maintained only if one makes a sharp distinction between imagining and conceiving; indeed, there are many more things one can imagine than one ever could conceive (this is the explanation of prophecy, see TTP i, III 15, 28; cf. my *Will of God*, Ch. 3). We can imagine a God-Lawgiver, but we cannot conceive a being that is infinitely perfect without seeing that it cannot be a lawgiver; we can imagine humans changing into animals, but we cannot conceive it, that is, we cannot conceive the chain of causes and effects leading to that event. Inversely, if one would have an infinite intellect (if one would be God), one would be unable to imagine anything (TIE § 54, II 20). Accordingly, the only worlds God can conceive are those entailed by the present world.

[41] CM I, 3, I 241. An idea of something impossible is itself impossible; Eth II, prop. 33, dem., II 116.

[42] A false idea is either a fictitious idea affirmed to be true or a contradictory idea, that is, a chimera or, in other words, no idea at all; TIE §§ 66–68, II 25–26.

[43] CM I, 1, I 233.

[44] CM II, 7, I 262.

rationis exist only in the mind is something they have in common with *entia ficta*, but there are at least two differences: (1) an *ens fictum* consists of at least two terms, whereas an *ens rationis* is simple; (2) an *ens fictum* excludes a clear and distinct perception, whereas an *ens rationis* is clearly and distinctly perceived, provided we do *not* believe that it is an *ens reale*.[45] Accordingly, the correct use of *entia rationis* is contingent upon their being "ideas of ideas", namely, upon the realization that they are *modi cogitandi*.

The background of this is undoubtedly formed by articles 48–50 of part I of Descartes' *Principles*. In these paragraphs, Descartes sets out to enumerate "all the simple notions which are the basic components of our thoughts."[46] According to him, "all the objects of our perceptions we regard either as things, or affections of things, or else as eternal truths which have no existence outside our thought."[47] Descartes does not attempt to enumerate the eternal truths because "we cannot fail to know them when the occasion for thinking about them arises, provided that we are not blinded by preconceived opinions."[48] However, among things and properties of things "the most general items are substance, duration, order, number, and any other items of this kind which extend to all classes of things."[49] Ultimately, however, there is only one fundamental division of things: they are either thinking things, pertaining to minds, or material things, pertaining to extended substance.

Spinoza obviously wants to change the ontological status Descartes assigns to "items [. . .] which extend to all classes of things," even if he specifically singles out 'substance', which of course he desperately needs for the *Ethics*.[50] In his eyes, all these items are neither things nor properties of things but *entia rationis*, that is, *modi cogitandi*. But this does not mean that they should be despised. A typical example is 'time'.[51] Outside the mind there is no time but only finite things with a certain duration, 'duration' being the existence of finite things (as 'eternity' is the existence of infinite things).[52] However, we can measure duration by means of 'time', that is, by devising the *ens rationis* called "time". So time is not identical to duration.

---

[45] That an *ens fictum* excludes a clear and distinct perception means, I presume, that if we imagine something (a thing or an event), this is properly imagined only to the extent that we have no clear view of the chain of causes and effects of which it should be a part – if we had such a view, the thing imagined would turn out to be either real or impossible; cf. TIE § 62, II 24. The difference with memory is that in this case the image of the object is accompanied with a specific notion either of its duration or of its place in time; cf. TIE § 83 (with note), II 31.

[46] *Principia* I, art. 47, AT VIII-A, 22.

[47] *Principia* I, art. 48, AT VIII-A, 22.

[48] *Principia* I, art. 49, AT VIII-A, 24.

[49] *Principia* I, art. 48, AT VIII-A, 22–3.

[50] CM I, 3, I 240; cf. Ep 12 (to Meijer), IV 54. Cartesian critics of Spinoza would not fail to concentrate on that very point; see Theo Verbeek, "Wittich's criticism of Spinoza" in Tad Schmalz (ed.), *Receptions of Descartes: Cartesianism and anti-Cartesianism in early modern Europe* (London: Routledge/Taylor and Francis, 2005), 113–27.

[51] For the following see CM I, 4, I 244.

[52] Cf. CM II, 1, I 250; II, 10, I 269; TIE §55, II 20; Eth I, def. 7, II 46; Ep 12 (to Meijer), IV 53–8, etc.

Duration is the existence of a finite thing, whereas time is a means to compare finite things to their duration. Accordingly, time is infinitely divisible (it must allow us to 'measure' the various durations of an infinite number of finite things), whereas duration, as the existence of a single finite thing, is indivisible (we cannot conceive half the existence of a thing, although we do say that the duration of one thing is twice the duration of another thing). There is nothing wrong in using the notion of 'time'. It becomes problematic only if we see it as a real thing or a real property; not as long as we see it as a *modus cogitandi*. What is true of time is also true of other *entia rationis*, like *opposition*, *order*, *agreement*, etc. All these "are perceived by us clearly enough as long as we think of them, not as things other than the essence of the things that are opposed, ordered, etc., but only as *modi cogitandi*, by which we retain or imagine things more easily."[53]

This makes it possible to clarify a few things. Notions like time, measure, and number undoubtedly fit Spinoza's description of the things referred to in proposition 39 as "whatever is common to human bodies and certain external bodies, which are common and proper to each whole as well as its parts." If we locate a body in time or count and classify bodies, we take them as wholes, assigning to each whole a certain duration, a certain place, and a certain position, which are also the duration, the place, and the position of their parts. Moreover, by means of the notions of time, measure, and number we create a common framework, not just for physics, which deals with the fundamental properties of matter, but also for the study of the biological, sociological, and moral world. Finally, this framework is ultimately centred in our own body, that is, in its relation to the rest of the world as it is experienced by our mind. In sum, the general purpose of propositions 38–40 is to suggest a strategy for the handling of *entia rationis*, that is, of all those principles, concepts, and notions which, without being ideas of *things*, are indispensable for the construction of a theory of reality. In a general way, therefore, reason can be said to be the faculty of devising and handling *entia rationis*.

At the same time, however, it should be pointed out that *entia rationis* are actually *entia imaginationis* and, by implication, that reason is closer to the imagination than to the intellect. This becomes clear from Spinoza's letter to Lodewijk Meijer on 20 April 1663.[54] According to that letter, time and measure are what we get if we detach duration and quantity from the things to which they belong; number is what we get if we detach the *modi* of substance from substance itself. So time, measure, and number "are *entia rationis*, that is, *auxilia imaginationis* (instruments of the imagination)." If we use these notions to understand true reality, especially if we do not understand the nature of these notions, we land ourselves into inextricable problems, given the fact that "there are numerous things which cannot be understood by the imagination but only by the intellect, like substance, eternity, and other things." In fact, by means of *entia rationis* one cannot even understand *modi*, "for if we try to do this, we detach them from substance and from the way they follow from eternity."

---

[53] CM I, 5, I 245. The same is true of notions like *genus* and *species*; CM I, 1, I 235.
[54] Ep 12, IV 52–62.

So even if reason makes it possible to *know about* the real world, it does not allow us to *understand* it. Moreover, reason is much closer to the imagination than it is to the intellect – in fact, it is different from the imagination only to the extent that it is active, whereas the imagination is essentially passive.

The way Spinoza deduces reason in the *Ethics* is not very satisfactory. This may be an effect of the way the *Ethics* is organized. Whereas the road ("method") suggested in the *Treatise on the Improvement of the Intellect* is to use the idea of God as the "norm" of truth and falsehood, which in turn would allow the deduction of "common notions" as "all things happen in a uniform way," the author of the *Ethics* cannot but see the notion of truth and reason as part of a psychological and genetic theory of the mind. This leads him, not only to blur the distinction between imagination and intellect, but also to presuppose, rather than explain, that from passive the mind can become active.

Chapter 6
Natural Law and the Theory of Moral Obligation

Thomas Pink

Abstract The chapter explores theories of moral obligation from those of late scholastics such as Francisco Suarez and Gabriel Vasquez to those of Samuel Pufendorf and John Locke. The theories of Pufendorf and Locke are contrasted. Although these two theories appear similar, there is a profound difference between them. In Pufendorf as in a scholastic such as Suarez, practical reason is seen as involving two distinct kinds of justificatory force or modes of justificatory support, recommendation and demand; and moral obligation is identified, not as a reason-giving property of actions, but as one of these justificatory forces, the force of demand, a force that directly binds the will. Whereas in Locke there is only one justificatory force, that of recommendation; and moral obligation is no more than a reason-giving property, the property of being commanded by a punitive God, among the many that generate this force. In Locke as in subsequent English-language philosophy, moral obligation ceases to be a justificatory force that directly binds the will, and comes to be no more than a reason-giving property of the voluntary actions that the will causes and motivates. The chapter expresses doubts about whether this development has been a genuine conceptual advance, and explores the problems it raises.

Keywords Obligation · practical reason · natural law · belief-desire model · Suárez Pufendorf · Locke

6.1 Moral Obligation and Blame

What is a moral obligation? What distinguishes those standards which are obligatory from those which are not? There is one very natural answer. A standard that is obligatory is a standard breach of which is, not foolish or merely inept, but blameworthy. We should understand obligation in terms of its link with the special kind of criticism that meets breaches of obligation – the criticism that is blame.

T. Pink
Department of Philosophy, King's College London, Strand

The advantage of appealing to blame is that it does seem to pick out what is so special about moral obligation – namely its demanding nature. Obligations demand our compliance and do not merely advise it, so to speak, in that breach of the obligation is not merely foolish, but blameworthy. Blame is a distinctively condemnatory criticism; and that makes moral obligation a potentially condemnatory, and so demanding, kind of a standard.

But there is an obvious problem with such a blame-centred approach. For what is blame? A criticism, with a content: that the person blamed is morally responsible for having done wrong. But what is wrong-doing? Breach of obligation. We are, it seems, moving in a circle. The notion we are trying to explain defines and forms the content of the very criticism that we hoped to use to explain it. The appeal of a blame-centred theory of obligation is obvious enough; but so is the threat of circularity.

The threat of circularity may well encourage us to say something more about blame; to say something more to characterize blame other than just that blame is criticism for doing wrong – for breaching obligation. And there are many such theories about blame, some particularly well known to modern philosophy. And each of these theories implies a different view of moral obligation, the standard that is to be understood in terms of blame. Some philosophers, such as John Stuart Mill, have seen blame as a kind of sanction or punishment: and that makes moral obligations standards one is to be punished for breaching. Whereas David Hume thought that blame was simply a fairly intense negative evaluation – a form of contempt. So for Hume moral obligation is nothing more than a kind of evaluative standard – a standard which it is contemptible to breach. In what follows I want to look at a way of understanding blame that is rather different from these. This is the attempt to understand blame not simply as an evaluation or sanction, but as a distinctive form of rational criticism.

This particular version of the blame-centred approach to moral obligation has entirely disappeared, at any rate from modern English-language philosophy – but it was historically important. We find it amongst some members at least of the scholastic or Catholic natural law tradition. These thinkers took blame to be a distinctive form of criticism that is *rational*. Blame on this scholastic theory is a form of rational criticism in that to blame someone is to criticize them for disregarding a force of practical reason. And that makes moral obligatoriness a distinctive kind of rational or argumentative or justificatory force – that force asserted and defended by the special kind of rational criticism that is blame. And natural law, the pre-positive moral law, is a law that embodies this force of practical reason.

Not only do we find this *Force model* of moral obligatoriness in Catholic or scholastic natural law theory. It lived on into the succeeding school of so-called modern or Protestant natural law as well. So I shall examine the conception in the Jesuit writer Francisco Suarez; then in Samuel Pufendorf, a Protestant theorist of law and morals who develops the conception further in late seventeenth century Germany in his *De Iure Naturae et Gentium*; and then in John Locke, a thinker who, despite many superficial similarities to Pufendorf, clearly abandons the conception for something much closer to the approach to moral obligation of

modern English-language philosophy. The story I will tell will be partly about a change in psychological theory; partly about a deep tension within normative theory.

6.2 Obligatoriness as a Force of Practical Reason

First, let me say something more about what a Force model of moral obligatoriness involves. Moral obligations are cited in practical arguments – in arguments about what to do. That's why it seems so natural to take moral obligations to involve a kind of reason for doing what is obligatory. For are not arguments in the business of providing reasons for what the arguments support? So, for example, someone claims that you should help your friend. Why? The argument continues: you have an obligation to help; it would be wrong not to help her after all she has done for you. The appeal to the obligation and its ground seems to be an appeal to a kind of reason – a reason of a particularly demanding kind.

But you could see moral obligation as involving a reason for doing what is obligatory and still not espouse a Force model of moral obligation. To understand more of what a Force model involves we need to look in very general terms at the structure of practical justification or reason, and at what elements it involves.

Motivation or will   *Voluntary Action*

Intention to do A   Doing A

Justify with given force   ⟵   Features F, G... of doing A
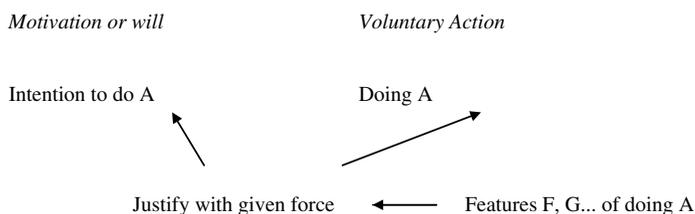
Consider actions that are *voluntary* – by which word "voluntary" I shall mean actions that we intentionally perform, when we do, on the basis of some prior motivating pro attitude towards performing the action, such as a prior decision or intention to perform it.

Deliberation or reasoning about how to act, practical deliberation, is principally and centrally about which such voluntary actions to perform – such as about whether to pay someone a sum of money, and the like. In deliberating we consider the various features which these possible voluntary actions have. We consider the actions both as possible ends in themselves, things possibly worth doing for their own sake, and as possible means to attaining further ends. Certain features of the voluntary – such as, say, the fact that paying someone money would fulfil a contract – then generate reasons or justifications for performing this voluntary action rather than another – reasons having a certain kind of force. What kind of force might this be? Not a causal force, still less one that just determines a response to it. For people can perfectly well disregard the force of reason, at the cost of counting as irrational for so doing. But even though no causal determination is involved, we still use the term "force" in this context to convey a kind of directive guidance: one that is not physical

but normative. That is, we use the term "force" to refer to the kind of support that a justification provides for what it justifies. We say of a consideration that gives strong support for a conclusion that it has "great force", and we talk of the "force" of an argument. And the nature of this justificatory force is clearly going to be linked to something else – the kind of criticism that someone is liable to if they disregard the justification provided. For, as we shall see, that criticism will assert and defend the force that has been disregarded.

It is important that we immediately respond to the force of this practical justification, when we do, by deciding or forming an intention to perform the action concerned. So the force of any practical justification will apply at two points: to support the voluntary action justified, but also, at the point of motivation or will, to support a prior decision or intention to perform that action. Take the fact that paying someone money would fulfil a contract. Suppose this feature of the voluntary action is a justification for performing it – for paying the money. If this justification is to move us to act, it cannot apply to the voluntary action of paying the money alone. It must also apply to the will or motivation on which the performance of that voluntary action depends. The justification must also support deciding to pay the money. Otherwise even a reasonable agent would note the justification for paying the money, but remain quite unmoved, as he would lack any justification for deciding or becoming motivated to pay it. And that would be absurd. Such justifications would no longer be practical. They would no longer move us to act.

There are then two aspects to any practical justification. First there is a feature of the voluntary from which the justification arises – a reason- or justification-generating feature, such as that to pay the money would fulfil a contract, a feature which gives us a reason to act. And such a feature may well be found only at the point of the voluntary action. It need not apply to the will as well. After all, it is paying the money and not intending to pay it which fulfils the contract. But then there is the force of the justification given – the kind of support which this feature gives to paying the money. And, if the justification is to move us to act, this force must apply, not just to the voluntary action, but to the will to perform it as well.

What kinds of justificatory force are there? One kind of justificatory force, certainly, is what I shall call the force of Recommendation. With this force reasons recommend the actions which they support, or make the performance of those actions advisable. In so doing they may defeat rival reasons for alternative actions – and thereby leave the action justified more advisable or sensible than those alternatives. To ignore the force of Recommendation is to be liable for criticism as foolish or as less than sensible.

The force of Recommendation is a force of reason – it is a force which reasons or justifications can certainly have. Indeed, I would go further, and conjecture that anything which counts as a reason will have some amount of this force. For to offer a reason for doing something is always, at the very least, to offer some recommendation for doing that thing, or to give some advice in its favour. The scholastics were I think perfectly familiar with this kind of justificatory force. It is the force of what they called *consilia* or counsels.

If we think that moral obligation is a kind of rational standard, we have to fit it into this general structure of practical justification. There is one obvious way of doing

so. This is to adopt what I call a *Feature model* of moral obligation. We simply model moral obligatoriness as a reason-giving feature. Suppose, for example, that I am under a moral obligation to pay you money because that way I would fulfil my contract with you. According to the Feature model, this means that because paying the money would fulfil a contract the action has a further feature – the feature of being morally obligatory. And this feature itself supports the action and moves us to act with the justificatory force which I have already mentioned as being essential to rationality – the force of Recommendation. The moral obligatoriness of paying the money makes it advisable or even more advisable to pay the money. Moral obligatoriness is a feature which adds to the advisability of doing what is obligatory. So Recommendation, this force that it is sensible to be moved by, less than sensible to disregard, is the justificatory force which all reason-giving features generate; and moral obligatoriness is one reason- or justification-giving feature that generates this force.

But there is another, quite different way of understanding the rationality of obligation. This is to adopt what I call a *Force model* of moral obligatoriness. According to the Force model, moral obligatoriness is not a further reason-giving feature at all. It is instead something quite different. It is a new and distinctive kind of justificatory force. In addition to the force of Recommendation, there is another force within practical reason – another way in which reasons can support actions. We might call this force the force not of Recommendation but of Demand. And according to the Force model, for an action to be morally obligatory is for it to be justified with this force of Demand. Suppose paying the money is morally obligatory because it would fulfil a contract. Then according to the Force model, the moral obligatoriness of the action consists in this: that the action would fulfil a contract is a reason for performing it, and a reason which does not merely recommend the action, or make its performance advisable, but demands its performance. And this different kind of justificatory support or force is evidenced by the distinctive kind of rational criticism with which disregard of it is met. Those who disregard the justification are not criticised as merely foolish or as less than sensible, but are blamed as wrongdoers.

6.3 Scholasticism and the Force Model of Moral Obligation

It is clear that a scholastic such as Suarez was committed to the Force model, not the Feature model. Besides *consilia*, there is in his view another *vis directiva* that guides our actions. This is the force of law, the obligatory demands of which are communicated not by *consilia* or counsels but by *praecepta* or precepts. And obligation under preceptive law clearly takes the form, not of any reason-giving feature but of a justificatory force. For Suarez makes an assumption about moral obligation which only the Force model validates. He regards it as a condition of moral obligatoriness moving us to act that it directly apply to the will itself, and not just to the voluntary action willed. If moral obligation is to move us to pay a sum of money, then, in Suarez's view, it follows that it is deciding to pay the money and not just paying the money that must be morally obligatory. Suarez was very clear about this, and regarded the point as quite uncontroversial in his tradition. The law of nature speaks

to us, he says, as the voice of our reason; and so it must apply to and direct the will
itself:

> So teaches Saint Thomas and on this point everyone. And the point is established because
> the law of nature is placed in reason, and immediately directs and governs the will. So it is
> on the will first and foremost that as it were by its very nature the obligation of the law is
> imposed. So the law is not kept unless through the exercise of the will.[1]

The assumption makes sense if moral obligatoriness is a kind of justificatory force.
That is exactly how the force of a justification moves us to perform a voluntary
action – by equally supporting the decision so to act. Notice that the assumption
would not hold at all were moral obligatoriness a reason-giving feature. Voluntary
actions have lots of reason-giving features that move us to perform them, but which
are not also found applying to the will. A voluntary action might be justified as
obeying a command, or as fulfilling a contract, or as helping the needy, without the
intention to perform the action counting as any of these things. Only the justificatory
support given by these features – the justificatory force which they exert – need
apply to decisions and intentions of the will.

With this force of moral obligation or Demand comes blame, as the distinctive
kind of rational criticism by which this particular force of reason is asserted and
defended. We find within the Catholic natural law tradition one important reading of
what blame is – an attempt to understand blame and its content more informatively
than by just blankly appealing to the mere idea of a wrong or breach of obligation.

The characterization of blame is evaluative. It sees blame as the imputation of an
action's value to its agent. The action's badness is imputed to its agent. This view of
blame is stated, magisterially and very influentially, by Aquinas:

> Hence a human action is worthy of praise or blame in so far as it is good or bad. For praise
> and blame is nothing other than for the goodness or badness of his action to be imputed
> to someone. Now an action is imputed to an agent when it is within his power, so that he
> has dominion over the act. But this is the case with all actions involving the will: for it is
> through the will that man has dominion over his action. [. . .] Hence it follows that good or
> bad in actions of the will alone justifies praise and blame; for in such actions badness, fault
> and blame come to one and the same.[2]

---

[1] "Modus voluntarie operandi, sub praeceptum legis naturalis cadit, et ideo necessarius est, ut lex
naturalis servetur. Ita docet D. Thomas q. 100 art 9 et ibi omnes. Et probatur, quia lex naturalis in
ratione posita est, et immediate dirigit et gubernat voluntatem; ergo illi imponitur quasi per se, et
principaliter obligatio illius legis: ergo non observatur illa lex nisi mediante voluntate: ergo modus
voluntarie operandi est per se praeceptus, ac necessarius ad observationem talis legis." (Francisco
Suarez, *Opera Omnia*, vol. 5: *De Legibus ac Legislatore Deo*, ed. Charles Berton [Paris: Vives,
1856–] 2.10 § 4, 123.)

[2] "Ergo actus humanus ex hoc, quod est bonus vel malus, habet rationem laudabilis vel culpabilis
[. . .] nihil enim est aliud laudari vel culpari quam imputari alicui malitiam vel bonitatem sui actus.
Tunc autem actus imputatur agenti quando est in potestate ipsius, ita quod habeat dominium sui
actus. Hoc autem est in omnibus actibus voluntariis: quia per voluntatem homo dominium sui
actus habet. [. . .] Unde relinquitur quod bonum vel malum in solis actibus voluntariis constituit
rationem laudis vel culpae; in quibus idem est malum, peccatum et culpa." (Thomas Aquinas,
Summa Theologiae [Turin: Marietti, 1950] 1, 2 q21 a 2, resp.)

It is very natural to understand blame as involving a negative evaluation of the person blamed. What else are we doing when we blame someone, but claiming that it was bad of them to do what they did? But notice one important further characteristic of Aquinas's understanding of blame. Blame asserts and assumes a special or moral responsibility – a responsibility that is for how we act. In blame the moral character of an action is imputed to its agent, and imputed to the agent because that agent himself determined that the action would be performed. The agent was the determinant of the action because he had a power over whether the action would be performed – the power of freedom or control of how one acts. That means that the force of obligation that blame asserts and defends is a justificatory force which is agency-specific. It applies to how we act – and to how we act in so far as our action is a locus of self-determining freedom. As Suarez put it, laws, standards with the force of obligation, govern only free, human action: *lex tantum datur de humanis actibus*.[3]

Notice again that since this justificatory force governs free action, leaving us free to disregard it and act wrongly and other than it justifies, the force is not that of efficient causation determining and necessitating us to act as it justifies. The *vis directiva* that is moral obligation is normative and not physical. It dictates how we should act, not how we actually will act.

Of course, if moral obligatoriness is a justificatory force that is agency-specific – that applies to our capacity for action and that alone – and if any justificatory force must apply to the will, that has an obvious implication. The will must itself be a capacity for action. And so Suarez believed, along with the rest of his tradition.

Our capacity for action was seen by the scholastic tradition as immediately exercised at the point of the will. Our freedom and capacity for deliberate action is first and foremost exercised, not simply in what I have called our voluntary actions – the actions such as paying money, or crossing the road or thinking about where to go on holiday, that we perform on the basis of deciding to perform them – but also in the prior decisions so to act. Besides the category of what they called imperated or commanded actions – actions involving capacities outside the will, and motivated by a prior will or decision to perform them – the scholastics also admitted a category of elicited actions of the will itself. And these were actions of *electio* and *intentio*, or deciding and forming intentions to act. It was in these elicited actions of the will that our agency began, and by which, through their effects, the rest of our agency was determined. For imperated actions occurred as efficiently caused effects of elicited actions of the will – effects that would be produced exactly as willed provided no intervening change of mind occurred, and no obstacle to acting as intended or as willed arose.

---

[3] Suarez is absolutely insistent that precepts of law only address free, and so perfectly voluntary acts: "Addo praeterea, loquendo de propria lege, de qua nunc agimus, tantum esse posse propter creaturam rationalem: nam lex non imponitur, nisi naturae liberae, nec habeat pro materia, nisi actus liberos [. . .]" (Suarez, *Opera Omnia*, vol. 5: *De Legibus ac Legislatore Deo*, 1.3 § 2, 7.)

The scholastic conception of moral obligation as an agency-specific justificatory force was (as I have discussed in more detail elsewhere[4]) linked to what I call a *practical reason-based* model of action – a theory of action as consisting in a distinctively practical exercise of our capacity for rationality, whereby we respond to and apply justifications for the voluntary, such as those provided with the agency-specific force of obligation. And we immediately respond to and apply such justifications at the point of the will, when we decided to perform this voluntary action rather than that. So action takes place primarily in elicited acts of the will itself – and then only secondarily in voluntary actions, the actions of other faculties that are caused by a prior will or decision and intention to perform them.

I have been presenting this late scholastic natural law theory as really a theory of a distinctive justificatory force within practical reason – a force that is agency-specific, just as is blame, the criticism with which it is linked. But was not the natural law also seen as a moral law in this sense, something that arises out of moral legislation – out of the decrees of God as moral law-giver? So it was by some, Suarez included. But the view of natural law as arising through divine legislation was not universal in late scholasticism. There were also those who saw the natural law as without legislative origin – as a law without a law-maker.[5] And Suarez, who did see the natural law as legislated, had to make sense of this position in terms of the more fundamental and shared view of moral obligatoriness as a justificatory force linked to blame. For Suarez moral obligation is still such a justificatory force; it is just that in his view, the feature of being decreed by a legislative superior such as God is necessary to generating this force – and so it is correspondingly necessary to the special badness that constitutes wrongdoing proper, the badness that is imputed in blame for doing wrong.[6] So the idea of natural law that was common property to Suarez and his contemporaries was not the idea of a set of divine decrees – the place of these in relation to natural law was an area of controversy – but the idea of something importantly different. It was the idea of something which generates a rational demand on action – a justificatory force that we are morally responsible for responding to.

---

[4] See my "Suarez, Hobbes and the Scholastic Tradition in Action Theory," in T. Pink and M. Stone (ed.), *The Will and Human Action: From Antiquity to the Present Day* (London: Routledge, 2003); "Action, Will and Law in Late Scholasticism," in J. Kraye and R. Saarinen (ed.), *Ethics on the Threshold of Modernity* (Dordrecht: Springer, 2004). Further discussion is to be found in my forthcoming *The Ethics of Action*, vol. 1: *Self-Determination* and vol. 2: *Normativity* (Oxford: Oxford University Press, forthcoming).

[5] A classic scholastic account of how natural law exists prior to and independent of any legislation is to be found in the work of Suarez's fellow Jesuit and argumentative opponent Gabriel Vasquez in his *In Primam Secundae Divi Thomae* (Ingolstadt, 1612), also discussed in my "Action, Will and Law in Late Scholasticism". For a clear indication of how representative Vasquez's views were, see that invaluable contemporary synopsis of late scholastic views on moral obligation, Poncius's supplement to the 1639 Lyon edition of Scotus's *Quaestiones in Librum Tertium Sententiarum*, Distinctio 37 (John Duns Scotus, *Opera Omnia*, vol. 7, ed. Luke Wadding [Lyon, 1639] 857–77).

[6] Suarez, *Opera Omnia*, vol. 5: *De Legibus ac Legislatore Deo*, 2.6 § 17, 110.

So we have an important package. We begin with a conception of moral obligation that is blame-centred – that understands it in terms of its link to blame. Blame is understood in turn as a distinctive kind of rational criticism, one by which the badness of an action is imputed to its agent, a rational criticism which is linked to an equally distinctive rational or justificatory force – one which is agency-specific, governing action as something which occurs at the point of the will, and which we have a special responsibility for as something that we determine for ourselves.

This scholastic natural law conception of obligation as an agency-specific justificatory force raised three questions. And these three questions are fundamental. It is these three questions, and how they were eventually to be addressed and resolved, which were to determine the conception's subsequent transformation and eventual fate.

First, how is this justificatory force a force of reason? We can certainly say, as the scholastics did, that reason contains *praecepta* as well as *consilia* – that alongside a force of counsel or recommendation it contains one of precept, of demand or obligation. But what makes this true? What makes blame, the criticism which is supposed to define Demand, a *rational* criticism? For how is the moral badness which blame asserts and imputes to the agent connected with irrationality? An action that is morally bad may count as such only because, as one late scholastic put it, it is inconsistent with rational nature.[7] But as yet it seems that the connexion between the action's badness and its irrationality is only being asserted, not shown. Wherein precisely lies the inconsistency?

Secondly, how are we to understand what moral badness is? Have we in the evaluative reading of blame an account of moral obligation that is genuinely informative; that avoids circularity?

Thirdly, we have the conception of agency as something exercised in decision making and intention-formation, at the point of the will – a will that is self-determined because free. This is the conception of agency that the theory demands. But is this conception of agency defensible?

6.4 Samuel Pufendorf's Force Model of Moral Obligation

The first two questions, the questions about practical reason and about the nature of blame, are central to the work of Samuel Pufendorf. Pufendorf is rightly understood as a theorist of the school of post-Grotian or modern or Protestant natural law. But it is important that he carefully develops and explores an account of moral obligatoriness that is in certain fundamentals just that of the Catholic school.

For Pufendorf, moral obligatoriness is what he calls an operative moral entity – a directive force that is imposed on us "chiefly for the direction of acts of the will"[8]

---

[7] See Vasquez, *In Primam Secundae Divi Thomae*, vol. 1, disp. 97.3, 659.

[8] "Nobis illud iam est dispiciendum, quomodo ad dirigendos voluntatis potissimum actus certum attributi genus rebus et motibus naturalibus sit superimpositum, ex quo peculiaris quaedam

through the decrees of God as our legislator. Pufendorf refers to moral obligations as laying "an inner bond as it were on the freedom of our will."[9] The function of their force is to direct the will as the immediate locus of self-determined freedom:

> There is one reason that man is capable of being subject to obligation; and this is that he has a will which can turn either way, and so can conform itself to a moral rule; it is quite otherwise with those beings who are intrinsically determined to one simple way of acting.[10]

Obligation binds the will morally, but it leaves it free to disregard this bond.[11] As with Suarez, the force of obligation is normative. It calls for a response, but does not necessitate it. And when disregarded this force of obligation is asserted and defended by blame – blame being taken to be a criticism in which the actions of our will are imputed to us as our responsibility, and as our responsibility because within our control:

> Now the reason why a moral action may belong and be imputed to someone [. . .] is none other than because it lay within his power and ability that such an action be done or not, be undertaken or omitted.[12]

In Pufendorf we have broadly the same view of obligation as a justificatory force as in the Catholic tradition, linked to the same view of action as occurring in and through the exercise of a free and self-determined will – as occurring in free elicited acts of the will itself, as well as in the further imperated actions of other faculties that decisions of the will motivate and explain. It might seem then that we have a theory of moral obligation just like that of the scholastic tradition – and particularly like that of a more voluntaristic member of that tradition such as Suarez, who also places the ultimate source of moral obligatory force in divine legislation.

Part though of what differentiates Pufendorf is his great sensitivity to two questions which did not loom so large for Suarez. There is the nature of blame itself. The scholastics saw blame as a kind of negative evaluation. But is this account of blame sufficiently informative? Does it do enough to explain the nature of obligation, the standard that blame asserts and defends? And then there is the place of

---

convenientia in actionibus humanis resultaret, et insignis quidam decor atque ordo vitam humanum exornaret. Et ista attributa vocantur entia *moralia*, quod ad ista exiguntur, et iisdem temperantur mores actionesque hominum, quo diversum ab horrida brutorum simplicitate habitum faciemque induant." (Samuel Pufendorf, *De Iure Naturae et Gentium* [Amsterdam, 1688], 1.1 § 2, 2.)

[9] "Ex hisce duobus fontibus manare arbitramur vim obligationum, quae intrinsecum velut vinculum voluntatis nostrae libertati injicere intelliguntur." (*Ibid*. 1.6 § 12, 70.)

[10] "Quod igitur homo aptus sit ad recipiendam obligationem, una quidem causa est, quia voluntatem habet, quae in utramque partem sese flectere, adeoque ad normam aliquam moralem se componere potest; secus atque illa entia, quae ad uniformem agendi modum intrinsece sunt determinata." (*Ibid*. 1.6 § 6, 63.)

[11] "Iniicitur enim per eandem [obligationem] velut fraenum aliquod morale nostrae agendi libertati, ut in diversam, quam quo illa ducit, partem tendere recte non possimus. Etsi per eam voluntas hautquidquam ita stringi queat, ut saltem de facto et suo periculo in diversa abire non detur." (*Ibid*. 1.6 § 5, 62.)

[12] "Caeterum quod actio moralis ad aliquem pertinere, eique imputari posit [. . .] eius causa nulla est alia, quam quod in potestate et facultate alicuius fuit, illam fieri vel non fieri, suscipi vel omitti." (*Ibid*. 1.5 § 5, 48.)

moral obligation within practical reason. Is moral obligatoriness really a force of *argument*, a justificatory force apt to move reasonable beings? What is the place of moral obligatoriness, what I have called the force of Demand, within wider practical reason? As we shall see, for Pufendorf these two questions are closely related.

Pufendorf's theory of blame is importantly different from the one suggested by Aquinas. It retains the theory that in blame the agent's wrongful action is imputed to him. But it is no longer the evaluative theory that Aquinas deployed.

Pufendorf is deeply unhappy with the idea that moral obligation can be understood in terms of the badness of actions which breach it. For what is this badness? It is clearly not *natural badness*, as Pufendorf terms it, or the badness that involves harming or worsening the interests of some agent or agents. For we can perfectly well have natural badness in an action without breach of obligation – as in animal actions.

> Many acts which tend to the welfare of men have the same natural effect among beasts, although among the latter they possess no moral quality.[13]

The kind of badness in an action that goes with breach of obligation and that blame asserts or implies is different. It is a specifically moral badness. But what then is moral badness? It seems just to be that kind of badness which we get in an action which is not merely harmful, but which breaches moral obligation – which is bad morally because morally wrong. But then we are arguing in a circle – and that is Pufendorf's complaint:

> If the definition of natural law is built upon that foundation of the necessary moral goodness or moral badness of some actions, it cannot avoid obscurity and circularity [...] in any definition of natural law, 'good' must be understood as natural good and not moral good, since, indeed, it would be absurd to define a thing in terms which presuppose that the thing defined is already known.[14]

The second problem is how to locate the justificatory force that is obligatoriness within reason. For Pufendorf is very clear that moral obligatoriness is a force that is rational to this degree – it is asserted in argument or reasoning about what to do, and which we respond to as reasonable beings. But that means that the force of moral obligation cannot just float free of what we independently understand practical

---

[13] "Verum haec ipsa bonitas et malitia naturalis actionem per se hautquidquam in genere morum constituit. Nam et multa sunt ad felicitatem et commodum hominis aliquid facientia, quae non sunt moraliter bona, utpote nec voluntariae actiones, nec ulla lege imperatae; et multi actus, qui in hominis commodum cedunt, eundem quoque in brutis effectum naturalem producunt, qui in his tamen nullam moralitatem habent. Sic a mutuis laesionibus abstinere, cibi et potu moderate uti, sobolis curam agere, ad brutorum conservationem aeque ac hominum faciunt; neque tamen ideo bruta actiones moraliter bonas exercere dicuntur." (*Ibid.* 1.2 § 6, 19.)

[14] Est praeterea et hoc monendum, si definitio iuris naturalis isti fundamento de necessaria honestate et turpitudine quorundam actuum superstruatur, eam non posse non fieri obscuram, et velut in circulum revolvi [...] in definitione legis naturalis bonum sumendum esse pro bono naturali, non morali; quippe cum definitio absurde fiat per ea, quae definitum iam ante notum supponunt. (*Ibid*. 2. 3 § 4, 126.)

reason to involve. The force of moral obligation must be rationally detectable, and it must be one that it would clearly be unreasonable to disregard.

As reasonable beings and even prior to any thought of law we are moved by considerations of natural good and bad – by considerations of benefit and harm to ourselves and to other humans. Pufendorf insists that the force of moral obligation must be anchored in such considerations of natural good and bad.

> [The natural law] so harmonises with the natural and social nature of man, that the human race can have no moral and peaceful social existence without it; or, if you prefer, it has a natural goodness and utility from its own native efficacy for the human race in general.[15]

And this is essential both to the knowability of what natural law requires, and to its argumentative force, its rational authority for us:

> We call an action good morally, or in moral estimation, which agrees with law; that action evil, which does not agree with it. For natural and material goodness, by which a thing or action is understood to tend to a man's convenience or improvement, is discussed in another place; although natural good is connected with the above moral good in those things which are enjoined by natural law, and usually by civil law as well, and gives that moral good some handle as it were among rational creatures.[16]

And so we arrive at the Pufendorfian account of blame and moral obligatoriness – an account that is designed to address both these problems, the problems of the content of blame and the rationality of complying with moral obligations.

What is the content of this censure that is blame? Blame is not simply an assertion that for what we have done we count as morally bad. As we have seen, for Pufendorf, our moral badness in fact reflects and presupposes the wrongness of our action, and does not explain it. Rather the content of blame is that we have done something for which we deserve some evil:

> Now many other things influence the will to turn to one side rather than the other. But an obligation has this difference from them: they bear down the will as by some natural weight, and on their removal the will returns of itself to its former indifference; while an obligation affects the will morally, and puts the will into such a frame of mind, that it is forced of itself to weigh its own actions, and to judge itself worthy of some evil, unless it conforms to the prescribed rule.[17]

---

[15] Illa est, quae cum rationali et sociali natura hominis ita congruit, ut humano generi honesta et pacifica societas citra eandem constare nequeat; aut, si ita mavis, quae bonitatem velut naturalem seu utilitatem ex nativa sua efficacia ordine ad genus humanum in universum habet. (*Ibid.* 1.6 § 18, 77.)

[16] "Actionem bonam moraliter, seu in genere morum, (nam de bonitate velut naturali et materiali, per quam res aut actio in commodum et perfectionem alicuius cedere intelligitur, alibi agitur, etsi haec circa ea, quae lege naturali, et ut plurimum etiam civili lege praecipiuntur, cum illa morali sit coniuncta, et inter animalia rationalia illi ansam velut dederit) dicimus, quae cum lege congruit." (*Ibid.* 1.7 § 3, 79.)

[17] "Porro licet multa alia voluntati, ut in alterutram partem vergat, momentum adferant; hoc tamen obligatio prae istis habet peculiare, quod ea velut naturali aliquo pondere voluntatem premant, quo remoto haec ad indifferantiam suam ultro redeat: obligatio vero moraliter voluntatem afficiat, et peculiari quasi sensu eandem intrinsece imbuat, ut in suas actiones ipsa censuram exercere cogatur,

As wrong-doers, we deserve some evil because we have freely disregarded and disobeyed the will of God as a legislator over us, a God who has also ensured that the evil we deserve we will in all likelihood meet. In decreeing the law, God has also willed the threat of a sanction that punishes breach of the law. And this is a sanction that it would be reasonable anyway to avoid, whether deserved or not. So the law is like a threat delivered against those who break it. But not just a threat is involved. The law also generates a distinctive and justificatory force of its own, the force of obligation – the force which is invoked in blame, by the thought that the evil threatened is also deserved:

> Again an obligation differs in a special way from coercion, in that, while both ultimately point out some object of fear, the latter only strikes the will with an external force, and impels it to choose some undesired object only by the sense of threatening evil; while an obligation in addition forces a man to acknowledge of himself that the evil, which has been pointed out to the person who deviates from an announced rule, falls upon him justly, since he might of himself have avoided it, had he followed that rule.[18]

Morally bad people may be moved to meet obligations by fear of sanctions alone. But the morally good will be moved by the force of obligation as well. So when the law of nature requires a person to seek peace and avoid conflict, there are two considerations that should move him. He should be dissuaded from a war of all against all by reason using "a twofold proof" – that to wage war without provocation, is both improper and unprofitable.[19]

Sanction is an important and fundamental part of Pufendorf's theory of natural law. The legislation which imposes obligations must always be accompanied by sanctions to enforce it. These sanctions can take two forms – natural and supernatural. The supernatural sanctions are those of the next life. But these sanctions cannot be known by reason to apply. Certainty as regards the threat of these depends on revelation. The natural sanctions, on the other hand, are the evil consequences which wrong-doing is likely to have for human nature in this life, and specifically for the wrong-doer himself. For Pufendorf is sure that the wrongdoer is always more likely than not to lose by his wrongdoing, and he will only ever get away with it in this life through luck. The natural risk of loss through ill-doing is still seen by Pufendorf as a legislatively imposed threat of sanction, because the creation by God of human nature with its understanding of reason and law and its susceptibility to such bad consequences is seen by Pufendorf as a legislative act. By so creating us with our

---

ac malo seipsam dignam iudicare, ni praescriptae sese normae reddiderit conformem." (*Ibid.* 1.6 § 5, 63.)

[18] "Et in hoc praecipue obligatio differt a coactione, quod licet utraque ultimo metum aliquem ostentet; haec tamen extrinsecus tantum voluntatem concutiat, et in electionem rei ingratae per solum sensum imminentis mali impellat: obligatio autem insuper efficiat, ut quis ipse adigatur agnoscere, sibi non immerito accidere malum, quod a proposita regula devianti est propositum, cum ultro eandem sequenti istud declinare licuerit." (*Ibid.* 1.6 § 5, 63.)

[19] "a bello, in quod per pravos affectus empellitur, quale est etiam illud quod fingitur omnium in omnes, per rationem duplici potissimum argumento revocatur; nimirum quod deprehendit, bellum altero non lacessente susceptum esse et indecorum et inutile." (*Ibid.* 2.2 § 9, 118.)

human nature, God has willed our conformity to the law that requires us to act in the ways generally essential to the good of that nature, and to avoid acting in ways likely to be harmful; and this is something which can be rationally determined, given the provability of the existence of God as the creative Providence responsible for giving us our human nature,

> For, indeed, the dictate of reason teaches us not only that the observance of natural law is profitable to human race, but that God also wills and commands mortals to guide their actions by the rule of these laws; and this is enough for the essence of law.[20]

The appeal to sanctions is doing three jobs. First, it helps mark out legislative acts – those acts by which the force of moral obligation is generated. Then, it is part of a system of natural consequences that link conformity to the law with natural benefits, non-conformity to the law to natural harms. And these links are essential to establishing moral obligatoriness as a force that directs us as reasonable beings, being one the application of which we can rationally detect, and which it must be unreasonable to disregard. And finally appeal to the desert of sanctions enables us to explain the content of blame in a way that tries to avoid the circularity problem that generally threatens blame-centred theories of obligation – and specifically the scholastic evaluative theory.

6.5 John Locke's Feature Model of Moral Obligation

It is worth briefly comparing Pufendorf's account of moral obligation with that in Locke's *Essay*. For the two, though superficially similar, are profoundly different. Superficially, Locke seems to have a view of moral obligation that is broadly the same. In Locke too we find moral obligation or duty as a standard on action – a standard which is imposed in a divine law by God as its legislator, and which is backed by divinely imposed sanctions:

> But what Duty is, cannot be understood without a Law; nor a Law be known, or supposed without a Law-maker, or without Reward and Punishment [. . .].[21]

But consider this vital difference. Pufendorf has an account of what kind of thing moral obligatoriness is that is in the same family as that of scholasticism. Whereas Locke's is fundamentally different.

For Pufendorf obligation is something distinct from legal decrees or commands. These decrees or commands are reason-giving features – whereas moral obligatoriness is a justificatory force which those decrees or commands generate – a force

---

[20] "Sane enim ex dictamine rationis non intelligitur solum, legum naturalium observantiam esse generi humano proficuam, sed et Deum velle ac iubere, ut mortales ad istarum normam actiones suas dirigant. Id quod ad essentiam legis sufficit." (*Ibid*. 1.6 § 4, 62.)

[21] John Locke, *Essay Concerning Human Understanding*, ed. P.H. Nidditch (Oxford: Clarendon Press, 1975), 1.3, § 12, 74.

with which these decrees and commands give us reason to comply.[22] And this justificatory force applies, like any such, not just to the voluntary actions which are commanded or decreed, but also to the will. This is plainly the traditional natural law conception of moral obligatoriness as an agency-specific justificatory force. On Pufendorf's view, then, it is not simply to the sanctions that accompany obligatory laws that we respond when we comply with moral obligations. We respond to the force of obligation itself – a force which presupposes those sanctions, but which is quite distinct from them.

While for Locke moral obligatoriness is not a justificatory force at all. It is merely the property of being legislatively decreed or commanded by God; and there is nothing more than that to our idea of moral duty or obligation:

> And thus we see, how Moral Beings and Notions, are founded on, and terminated in these simple Ideas, we have received from Sensation or Reflection [...] If I have the Will of a supreme, invisible Law-maker for my Rule: then, as I supposed the Action commanded, or forbidden by God, I call it Good or Evil, Sin or Duty [...] So that whencesoever we take the Rule of Moral Actions; or by what Standard soever we frame in our Minds the Ideas of Vertues or Vices, they consist only, and are made up of Collections of simple Ideas, which we originally received from Sense or Reflection: and their Rectitude or Obliquity, consists in the Agreement, or Disagreement, with those Patterns prescribed by some Law.[23]

So moral obligatoriness is a property just of the voluntary actions commanded by God – it is just that property of being commanded – and is not a force directly binding the will itself; and what we respond to when we comply with the obligation is simply the threat of sanctions that accompany these obligation-constitutive decrees, and so to the prudential advisability of avoiding those sanctions:

> Moral Laws are set up as a curb and restraint to those exorbitant Desires, which they cannot be but by Rewards and Punishments, that will over-balance the satisfaction any one shall propose to himself in the breach of the Law.[24]

What has happened is an important change. The justificatory force of Demand has disappeared entirely. All that is left is the feature that on Pufendorf's theory generated it, the feature of being subject to sanction-backed decree or command. And moral obligatoriness has moved from being a distinctive kind of force with which other features of an action may justify its performance to being just one further such reason-giving feature. We have moved from a Force model of moral obligatoriness to a Feature model. And this is a quite fundamental change in the theory of normativity, both of moral obligation in particular and of practical reason in general. Moral obligatoriness has wholly changed its metaphysical category, and the theory

---

[22] Pufendorf is crystal clear on this point. In book 1 chapter 6 in paragraph 4 he first defines law as a decree by which a superior obligates his subject. Then in paragraph 5 he defines an obligation as something different – an operative moral quality in the subject generated by the legal decree, a moral quality which binds the will.

[23] Locke, *Essay*, 2.28, § 14, 358.

[24] *Ibid*. 1.3, § 13, 75.

of practical reason has become drastically simplified. Practical reason has entirely lost the force of Demand, and now contains only that of Recommendation.

6.6 Psychology and Normativity

One factor that may foster or reinforce such a change has to do with moral psychology. I have described this factor in some detail elsewhere,[25] but let me say something about it here. It is a change in the theory of action, and one which proves very important within the English-language tradition; Pufendorf, by contrast, fiercely resists it.[26] This change in action theory is pioneered by Hobbes.[27] This is the restriction of free action to the category of voluntary or willed action – to the category of what the scholastics called imperated action. The category of pre-voluntary elicited action is abolished. This makes quite impossible the theory of obligatoriness as an agency-specific force that was common property both to Pufendorf and to the scholastic or Catholic natural law tradition.

Suppose free action can only occur in the form of voluntary action – action explained by a prior will to perform it. Then the prior exercise of the will itself must now be something that generally precedes free action, rather than providing its primary instance. Which means in turn that if obligatoriness is to remain a demanding standard that applies only to free action, it must now apply only to voluntary action, and not to prior non-voluntary motivations of the will. But as we have seen, any justificatory force must apply to these non-voluntary motivations, as well as to our voluntary actions. In which case, as restricted to governing voluntary actions and voluntary actions alone, moral obligatoriness can no longer be a justificatory force at all. It must instead be something else – something that can apply to the voluntary alone. It must be a reason-giving feature, such as the feature of being commanded.

But it is not merely change in the theory of action that is likely to foster the new approach to obligation. There is a problem elsewhere as well. The Pufendorfian account of how the force of Demand is a force of reason, a force that moves us as reasoning beings, looks profoundly unstable. If moral obligation is a force of reason, then there must be something about morally obligatory actions which moves reasonable beings to perform them. Pufendorf regards this as established by something he regards as essential to moral obligation – the naturally good consequences that follow conformity to natural law, and especially the naturally bad consequences that follow its breach, including those that are likely to fall as sanctions on the wrong-doer. But leaving things like that, it seems that the rational support for performing an action comes just from these natural consequences and sanctions,

---

[25] See my, "Action, Will and Law in Late Scholasticism".

[26] See Pufendorf's exposition of a broadly scholastic account of human freedom as exercised primarily in elicited acts of the will itself in book 1 chapter 4, and his animadversions there against Hobbes.

[27] See Hobbes in *The Questions Concerning Liberty, Necessity and Chance, clearly stated between Dr Bramhall Bishop of Derry, and Thomas Hobbes of Malmesbury* (London, 1656).

and the recommendation which those features anyway give of the action. Pufendorf may treat the action's obligatoriness as if it were a further force of reason – he terms it the force of a further rational argument or proof besides the appeal to natural good or evil – but nothing has been done to show that the obligation itself involves any further rational force at all. Such connexion with reason as has been established has only come through the natural consequences for good or evil of conforming to or breaching the law, and the Recommendatory force which they anyway carry. In which case, then, as far as reason is concerned, it still seems to be these consequences and the recommendatory force which they generate which is really providing the argument. If there is anything more to the specifically rational case for performing the action, this has not in fact yet been shown.

In which case why not abolish the force of Demand as an idle wheel, and substitute for it simply the feature of being subject to sanction-backed command? Instead of identifying moral obligatoriness with a mysterious second force of practical reason, we identify it with a thoroughly familiar kind of reason-giving feature.

For we already have one conception of obligation, the conception of a *legal obligation*, which works in just this way. For what else is being legally obligatory – obligatory under positive law – but this: the feature of being subject to a sanction-backed legislated decree or command, a feature which may well give us some reason to comply with the command.

The Feature model of moral obligation is now largely dominant within the English language tradition – even if views have changed as to what kind of reason- or justification-giving feature moral obligatoriness is. Thus T.M. Scanlon in our day certainly does not think that moral obligatoriness is the feature of being subject to sanction-backed divine decree. But he still thinks that moral obligatoriness has to be conceptualized as some kind of reason-giving feature. His argument for what in his view that feature is – the feature of an action's being part of any reasonable social contract – openly appeals to the Feature model. We should, he claims, understand moral obligatoriness in contractual terms, because that way we can best explain how an action's moral obligatoriness does give us so much reason to perform it.[28]

By contrast, if we accept the Force model, there is no question of determining how much reason moral obligatoriness gives us for doing what is obligatory. On the Force model, moral obligatoriness is no more a reason-giving feature than is its correlate advisability. If the Force model is correct, then, just like its advisability, an action's moral obligatoriness merely reflects the presence of other reason-giving features – in this case features that demand the action's performance and do not merely recommend it. Moral obligatoriness is just the force with which those other features give reason to perform the action.

---

[28] See the discussion in T.M. Scanlon, *What We Owe to Each Other* (Cambridge: Harvard University Press, 1998), 149ff. In Scanlon's view, moral obligatoriness is supposed to be a reason-giving feature of action; and it is reason-giving because an action's moral obligatoriness consists in the fact that reasonable social contractors would always agree that such an action is to be performed – and this fact is alleged to give each of us reason to perform the action.

Pufendorf did attempt to relate moral obligatoriness to natural good and evil and to the recommendatory justifications which that natural good and evil provide. But this was not to show how much reason an action's moral obligatoriness gives us for performing it. Given his allegiance to the Force model, that would have been an incoherent thing for him to try to do. Rather Pufendorf was trying to establish that moral obligatoriness really is a genuinely justificatory force – a force of reason. And to do that he felt he had to show that morally obligatory options are also rationally recommendable. In other words, Pufendorf was working on the following very plausible assumption.

Basic to practical reason is the idea of advice. To argue or reason for anything is always to give advice in its favour; it is always to make some appeal to the force of Recommendation, a force that had better not be defeated if the argument or reasoning is to be good. So if there is any other kind of justificatory force within practical reason, such as a force of obligation or Demand, then this force really must be consistent with the force of Recommendation. The options that the force of Demand supports must always be recommendable too. The obligatory must always be the advisable as well – something that is anyway a good idea.

But then we run an evident risk. The risk of course is that the justificatory work is effectively being done at the level of Recommendation; that in Recommendation we already have all the justificatory force we need to explain why reasonable people are led to do what is obligatory. By thoroughly integrating the theory of practical demand into a prior theory of reason as based on natural goodness, a theory in effect of Recommendation, as Pufendorf sought to do, we risk allowing the force of Demand to disappear. And so, for modern philosophy, it has effectively been allowed to disappear, and with it the obvious burdens which this force of Demand places on the moral psychology of action and the will.[29]

---

[29] An account of moral obligation which defends the Force model and which addresses these difficulties is to be found in my "Moral Obligation," in A. O'Hear (ed.), *Modern Moral Philosophy* (Cambridge: Cambridge University Press, 2004); and as part of a general account of normativity in my forthcoming *The Ethics of Action: Normativity*.

Chapter 7
Aspects of Inductivism in Thomas Reid's Science of the Mind

Riku Juti

Abstract The chapter deals with the nature of Thomas Reid's inductivism, its place in his epistemology and philosophy of the human mind, and its role in the methodological turn of British thought. It is argued that Reid's concept of induction is Newtonian, i.e. induction is part of the classical method of analysis and should therefore be understood as an instrument of discovery rather than of justification. Contrary to the received view, Reid does not reject the use of hypotheses in science. According to Reid, legitimate hypotheses are well-staged predictions that are intimately linked with the experimental setup at hand. As such, Reid's methodology keeps away from any hypostasis of spatial elements in cognitive activity. According to him, "immediacy" and "directness", which both characterize the activities of the mind, do not mean "presence". The distinction between things in the mind and things external is not meant to signify the place of things we speak of, but their subject. The chapter ends with an analysis of Reid's famous principles of common sense. It is suggested that these principles could be seen as related to medieval thinkers' "general propositions". Reid was an ontological nominalist. He needed nontrivial initial premises to back up his particular observations and facts in order to reach universal or general conclusions about nature.

The late eighteenth-century Scottish philosopher Thomas Reid (1710–1796) is known for his powerful attack on David Hume's sceptical speculations and for his advocacy of common sense as a source of knowledge. Closely related to these issues and supporting them was his famous rejection of the representational theory of ideas, which in Reid's view, was not only a false theory of concept formation but also the main cause of Berkeley's idealism, and consequently, also of Hume's scepticism – the former being, according to Reid, a half-way house of the latter.

R. Juti
Department of Philosophy, University of Helsinki

Several important studies have been written on these epistemological issues. Yet these issues are only lines of argumentation in Reid's overall philosophy, which mainly consists of two intellectual motifs – one being his advocacy of what he thought to be the philosophical content of Newton's methodological ideas, i.e. the promotion of induction in the logic of science; the other his desire to become – to use a well-known slogan – the "Newton of the mind", i.e. his focus on creating a new science of the mind, or "pneumatology", as he called it.

These two "leitmotivs" – Reid's dedicated inductivism and its role in the methodological turn of British thought and his aspiration to found a new science of the human mind – are also well-covered in the literature.[1] What is less known, however, is how they are interconnected and how they influenced Reid's anti-sceptical epistemology in general. As we do not know much of these interconnections, it is fair to say that we still do not have a comprehensive picture of Reid as a methodologist, scientist, and epistemologist. My aim in this chapter is to study Reid's inductivist methodology together with his scientific mental philosophy and, to that extent, contribute to a more balanced and thorough account of Reid's philosophy.

7.1 Epistemology vs. Mind Science

Reid's epistemology continues to be a popular topic of study among historically oriented philosophers and a source of inspiration for epistemologists who often sympathize with its pioneering externalism and naturalistic spirit. Unfortunately, Reid's theory of knowledge is sometimes interpreted as a kind of self-contained "intellectual entity" comparable in its scope and content with fully developed epistemological theories of contemporary philosophy. Such a comparison is bound to miss the real significance of Reid's epistemological pursuits.

First, Reid never worked exclusively on epistemology, as we understand that subject today. He was not in the analytical business of trying to define what knowledge is or what conceptual criteria can be given to the epistemic justification of beliefs. This *narrow* sense of epistemology was quite alien to Reid. Secondly, he never tried to epistemologize philosophy the way some of his predecessors (Berkeley, Hume, Kant) had done,[2] nor did he ever think that epistemology in the narrow sense – either in its transcendental or normative version – should take any leading role in the philosophical explanation of man's cognitive capacities.

Studying and viewing Reid's epistemology in isolation of his scientific methodology on the one hand, and of his scientific philosophy of the mind, on the other, tells more about us and our post-Kantian obsession with epistemology than about the authentic content of Reid's philosophy.

---

[1] Cf. e.g. Larry Laudan, *Science and Hypotheses* (Dordrecht: Reidel, 1981), 86–110.

[2] Cf. Hume's obsession with scepticism and Kant's efforts to overcome it leading to the Copernican revolution of philosophy.

The fact is that Reid uses the word "knowledge" surprisingly rarely in his writings (and he naturally does not use such expressions as "theory of knowledge" or "epistemology" at all),[3] and when he does use it, he deals with it in a fashion that is more psychological and cognitive than normative or conceptual. A good example is the following passage in *Essays on the Intellectual Powers of Man* where Reid defines knowledge as a category of judgment:

> In knowledge, we judge without doubting; in opinion, with some mixture of doubt [. . .] in knowledge [judgment] is more firm and steady, like a house founded upon a rock. In opinion it stands upon a weaker foundation, and is more liable to be shaken and overturned (IP, VI, III, 426a–b).

The purpose of this passage is to respond to Locke's theory, according to which knowledge is different from judgment, thereby making the faculty or concept of knowledge a more autonomous subject of study: the way we think of it today. Reid bluntly rejects Locke's theory and regards knowledge as a kind of judgment and judgment as a philosophically more fundamental concept than knowledge.

So although Reid is undeniably driven by some epistemological concerns, he seems to relate these concerns to the theory of the human mind that he is developing. One could almost say that he is interested in the metaphysical underpinnings of his epistemological ideas. For behind these ideas is his theory of the human mind – "pneumatology" – the aim of which is to transform the traditional metaphysical theory of the souls or spirits into a Newtonian science.[4]

It is true, of course, that not only Reid but also Descartes, Berkeley, Hume, and Kant (among others) were all interested in the metaphysical foundations of their epistemologies and linked their respective theories of knowledge to the theories of the human mind they were developing. None of them were in the analytical business of knowledge definition exclusively but rather saw their epistemological pursuits in the context of a wider inquiry into the conditions of knowledge. The point is, however, that while in Descartes, Berkeley, Hume, and Kant's case these epistemic conditions became formative of their ontological theories,[5] in Reid's case no such thing happened. In fact, it is debatable whether Reid addresses the question of the

---

[3] See, however, Reid's critical discussion of Hume's concept of human knowledge in *Essays on the Intellectual Powers of Man* (hereafter IP) Essay VII, Chapter IV, 484b–489b, in *Philosophical Works* with notes and supplementary dissertations by Sir William Hamilton, 2 vols., 8th ed. (Edinburgh: James Thin, 1895).

[4] As modern readers we are sometimes perplexed by the language used by the old masters, and are, perhaps, too quick to think that an odd choice of words must indicate a slip of the tongue on their part rather than a misunderstanding on our part. Reid writes, for instance, in a letter of 1763 to his famous opponent David Hume: "I shall always avow myself your disciple in metaphysics. I have learned more from your writings in this kind, than from all others put together." I think we can take Reid's reference to "metaphysics" in this letter quite seriously. I do not think that it is meant to be taken metaphorically. It is metaphysics – metaphysics of the mind – that Reid quite rightly thought he was pursuing in his work.

[5] In Descartes' case the first-person point of view is formative of the mind–body dualism, in Berkeley and Hume's case perception and imagination are formative of reality in general, and in Kant's case "metaphysics deals not with objects but with cognitions" as he says in his *Reflexionen*

conditions of knowledge as such at all. Reid's response to scepticism is to outflank it by introducing Newton's well-tested methodological ideas into the study of the intellectual powers of man. Epistemology as a conceptual enterprise and as a transcendental or normative study of the conditions of knowledge plays second fiddle to this Newtonian-*cum*-scientific approach to man. And as such, epistemology cannot have any formative role in the business of ontology, either.

By the same token, it is quite misleading to call Reid the "Scottish Kant".[6] This is misleading because Reid's philosophy did not promote the epistemological turn of western philosophy the way Kant's transcendental philosophy did. In Kant's philosophy epistemology actually transforms metaphysics. Ontological issues become epistemological. Kant explicitly says that metaphysics is not about objects but rather about reason and the structure of human cognition. When it comes to causation, for instance, Kant thinks that one should not investigate first the concept of cause but rather the faculty by which it is possible for us to have a priori causal knowledge. And investigating this faculty means investigating the elements of "pure thought", such as the forms of judgment.

Although Reid is interested in the structure of human cognition, he does not claim that questions of metaphysics could be reduced to the study of this structure. On the contrary, the human mind is, for Reid, a category of being among other categories of being, and as such an object of study like the rest of metaphysics. Reid does think that his science of the human mind transforms that part of classical metaphysics that focuses on spirits and souls, and that it does this in the same way as Newton's natural philosophy transformed another part of classical metaphysics. He also thinks that epistemological issues are among the issues addressed by his mind science, but he does not think that some epistemological principles – whether natural or transcendental – could transform metaphysics just by themselves.

Thus, the scientific revolution Reid is trying to initiate is really Newtonian rather than Copernican. He wants to change traditional metaphysics, but he wants to make it a Newtonian science rather than a part of epistemology. Epistemology becomes a part of that Newtonian science. Instead of explaining metaphysics either in the sense of laying down its normative foundations or in the sense of being in any way "formative" of it, epistemology is itself explained by a scientific theory of the intellectual powers of man that, according to Reid, replaces the traditional metaphysics of spirits and souls.

7.2 Mind, Science, and Inductivist Methodology

The relationship between Reid's methodological inductivism and his scientific philosophy of the mind has been overlooked in the literature. This is somewhat surprising considering the major role both these themes play in his philosophy and

---

on the *Critique of Pure Reason*, no. 91; also quoted in Ernst Cassirer, *Kant's Life and Thought* (New Haven and London: Yale University Press, 1981), 153.

[6] Cf. T.J. Sutton, "The Scottish Kant?" in M. Dalgarno and E. Matthews (ed.), *The Philosophy of Thomas Reid* (Dordrecht: Kluwer Academic Publishers, 1989), 159–92.

the obvious fact that they are interconnected. When this topic is touched upon, the discussion is normally reduced to the question of how Reid's anti-hypotheticalism relates to his rejection of the representational theory of ideas. Although there is a clear connection between these two issues in Reid's philosophy, that particular connection does not represent the most important effects of inductivism in Reid's philosophy of mind.

One reason why philosophers might have been reluctant to discuss these two themes (i.e. inductivism and mind science) together is the fact that while Reid is famous for his study of the human mind, he is, at the same time, rather infamous for supposedly supporting and promoting simple enumerative induction in scientific reasoning. Most present-day philosophers would say that sheer enumeration of particular facts and their inductive generalization gives no sufficient account of what really goes on in scientific thinking – Newtonian or any other type. Such induction tells very little about how scientific theories are verified and even less about how they are discovered. The problem with Reid is that he seems to think that enumerative induction tells all there is to scientific thinking. In particular, he seems to reject the use of hypotheses in science. Although he says in a famous passage of *Essays on the Intellectual Powers of Man*:

> Let hypotheses suggest experiments, or direct our inquiries; but let just induction alone govern our belief (IP, II, III, 251a).

This half-hearted recognition of the value of hypothesis as an investigative tool seems to be contradicted by other statements in which Reid clearly indicates that no useful discovery in the works of Nature has ever been made by way of hypothesis.[7]

Thus, Reid's open rejection of the method of hypothesis and his stubborn insistence on the supremacy of enumerative induction have made him – at least in some scholarly circles – a kind of intellectual villain. This has meant, among other things, that philosophers who sympathize with Reid's work tend to focus on his scientific philosophy of the mind and ignore his views on the logic of science, while those who relate to his work more critically like to highlight his insufficient methodology and ignore his achievements in mental philosophy. This is, of course, a serious simplification, but – as in most simplifications – there is a grain of truth in it.

---

[7] Such statements, or statements that clearly indicate Reid's open hostility towards hypotheses, can be found, for instance, in Chapter III "Of Hypotheses" of Essay I and Chapter III "Hypotheses concerning the Nerves and Brain" of Essay II of *Essays on the Intellectual Powers of Man*. Let me quote the following somewhat sarcastic passage from the latter of the two chapters mentioned above: "The votaries of hypotheses have often been challenged to shew one useful discovery in the works of Nature that was ever made in that way. If instances of this kind could be produced, we ought to conclude that Lord Bacon and Sir Isaac Newton have done great disservice to philosophy by what they have said against hypotheses. But, if no such instance can be produced, we must conclude, with those great men, that every system which pretends to account for the phenomena of Nature by hypotheses or conjecture, is spurious and illegitimate, and serves only to flatter the pride of man with a vain conceit of knowledge which he has not attained."

I think that Reid is neither a villain nor a hero. While his methodology is not, in fact, that bad, his scientific philosophy of the mind is either not that great. Most importantly, both must be seen in the context in which they were developed. In what follows I will discuss some strengths of Reid's inductivism – in particular, in relation to his philosophy of the mind – and also some weaknesses of his philosophy of the mind in relation to his inductivist methodology.

7.3 In Defence of Reid's Inductivism

Let us begin with Reid's inductivism. The received view of it can be summarized as follows. Scientific advances are made by patient observation based on accurate experiments and by conclusions induced from such observation. All theorizing, i.e. all hypotheses and conjectures that are not inductively reached and are not based on patient accumulation and enumeration of observational fact, is according to Reid, mere guessing and must, therefore, be excluded from science proper – in particular, from scientific belief formation.

Much has been written on the historical causes and effects of Reid's position – how it mirrors Newton's famous rules of philosophizing, how much of it derives from Lord Bacon, how much of it comes from his religious views, how it changed and charged the discussion on scientific method, etc. – and I do not propose to spend much time going over these connections. It suffices to say that although Larry Laudan is mainly right in arguing that Reid played a major role in the rise of the empiricist-inductivist tradition in philosophy of science, the real form and content of Reid's inductivist experimentalism as well as its proper place in the complicated history of scientific methodology is still widely unknown.[8]

I want to concentrate rather on some other issues that Laudan raises in his book *Science and Hypothesis*, in particular, on what he says about the topic of inductivism *vs*. hypotheticalism and the impact of Reid's philosophy on the methods of science. These are issues that also shape our judgement of Reid's methodology in relation to his philosophy of the mind. It is well known that Laudan criticizes those eighteenth-century empiricists who were opposed to the use of hypotheses and conjectures in science. In particular, he is critical of the official account of the emergence of the hypothetico-deductive method which focuses on the nineteenth-century geniuses Herschel and Whewell and ignores the pioneering spirit of the eighteenth-century hypotheticalists (Hartley, LeSage, and Lambert) "who between them elaborated the hypothetico-deductive method at a time when the major philosophers of the day [Hume, Reid, Condillac, Diderot, and Kant] had nothing to say in its favor."[9]

---

[8] Laudan, *Science and Hypotheses*, 86–110. Cf. also Robert Callergård, *An Essay on Thomas Reid's Philosophy of Science*, Acta Universitatis Stockholmiensis, Stockholm Studies in Philosophy 28 (2006) and my review of Callergård's work in *Theoria*, 4, 2007.
[9] Laudan, *Science and Hypotheses*, 17.

As Reid was one of the most influential inductivists who was particularly successful at bashing early proponents of hypotheticalism – such as David Hartley, for instance – he is understandably rather heavily criticized by Laudan for going too far in his rejection of the methodological novelties of the hypotheticalists and in his stubborn insistence on induction as a sole and sufficient method of proper science:

> But in a classic case of babies and bath-water, Reid's requirement [of eschewing hypotheses] went too far. His very narrow observational construal of the ground for warrantedly asserting the existence of a thing left Reid completely unable to give an account of the success of the many deep-structural theories of his time. By demanding too much, his epistemology was altogether unable to come to grips with contemporary theoretical sciences.[10]

Laudan may be right in his handling of Reid and other eighteenth-century inductivists in the sense that Reid and his fellow inductivists were largely understood as promoting a philosophy that seemed to undermine the importance of hypothetical theorizing in science. However, it is important to notice that Reid does allow a certain technical (i.e. investigative) role for hypotheses in scientific thinking and writes about some scientific hypotheses in a surprisingly positive light, for instance, about the hypothetical matter theory of Roger Boscovich. Moreover, he dedicates a full section of his *Inquiry* to "Squinting Considered Hypothetically".[11] This would be highly surprising if it came from a philosopher categorically opposed to hypothetical theorizing in science. The fact is, of course, that Reid is not opposed to such theorizing. Deep down his position is much more sophisticated than what Laudan seems to think.

First, Reid's construal of "the ground for warrantedly asserting the existence of a thing" is not as narrowly observational as Laudan suggests. It is important to understand that what Reid calls "observation" always takes place in an experimental setting. What are observed are not some simple, transparent, or self-intimating data ready for generalization. Although Reid says that "in the humble method of information, from the great volume of Nature we must receive all our knowledge of Nature," he does not mean to say that all information is as if *trivially* given to us in observation. Scientific genius is needed to set up "well-contrived experiments" where "Nature is put to the question". What we observe depends, therefore, on what we have decided to ask from Nature. It is in this interrogatory context that the scientist brings "his fine imaginations and most ingenious conjectures to the fiery trial of experiment and induction."[12]

Secondly, we do not really know what Reid would say about "warranted assertion". This is a type of epistemology Reid is not directly dealing with. He does say that "let just induction alone govern our belief," meaning that degrees of belief should reflect the amount of experimental evidence and that a firm belief – a

---

[10] *Ibid.*, 126–27.

[11] *Inquiry into the Human Mind* (hereafter **I**) Chapter VI, Section XV, in *Philosophical Works* with notes and supplementary dissertations by Sir William Hamilton, 2 vols., 8th ed. (Edinburgh: James Thin, 1895).

[12] All quotations of Reid in this paragraph come from IP, VI, VIII, 472b.

kind of end result – should not be formed solely on the basis of hypothesis. But, then again, he also says that "let [hypotheses] suggest experiments, or direct our inquiries." What is the epistemic status of such "suggestion or direction"? According to Reid, conjectures based on analogical reasoning "may have a considerable degree of probability, yet it is evidently in the nature of conjecture to be uncertain. In every case the assent ought to be proportioned to the evidence; for to believe firmly what has but a small degree of probability, is a manifest abuse of our understanding" (IP, I, III, 235a). Reid seems to think that the assertion of hypotheses (concerning the existence of a thing) is warranted only to the degree that allows the use of these hypotheses in directing the course of scientific discovery.

Thirdly, and most importantly, one should say in defence of Reid, *pace* Laudan, that the early proponents of hypotheticalism (e.g. LeSage, Hartley and Boscovich), however much they belong to the same tradition as the nineteenth-century geniuses Herschel and Whewell, still fully deserved the bashing given to them by Reid and his fellow empiricists. For it is obvious that the early hypotheticalists required too little of their hypotheses, and for reasons opposite to those of the empiricists, were at least equally unable to lay down the methodological foundations of the scientific theories of the early modern era.

Laudan admits this when he explains exactly what the weaknesses of the early hypotheticalists were. The latter were mainly interested in saving the known phenomena. They believed that hypothetical theories can be justified by the mere fact that they explain the very phenomena they were created to explain, i.e. that the fact that these theories explain these phenomena gives evidence for them. They never understood that this does not suffice for the purposes of justification; they did not understand that their way of describing the method of hypothesis made this very method totally arbitrary – they did not understand these issues even after Reid had told them that

> There never was an hypothesis invented by ingenious man which has not this [kind of] evidence in its favour. The vortices of Des Cartes, the sylphs and gnomes of Mr. Pope, serve to account for a great variety of phenomena (IP, II, III, 250b).

It is well known that it was only in the nineteenth century that hypotheticalists understood that a hypothesis must be successfully tested against a body of evidence independent of the circumstances which the hypothesis was invented to explain in the first place, i.e. that it must be tested against the unknown rather than the known. It is only when a hypothesis is able to predict states of affairs previously unknown that it will have real evidence in its favour, and it is only through such tests against the unobserved and the unknown that the method of hypothesis is freed from arbitrariness.

But, according to Laudan, Reid is equally, or even more, to blame because of his narrow-minded inductivism. If the early hypotheticalist failed to follow the basic guidelines of the method of hypothesis, Reid failed to perceive its true prospects.[13] Furthermore, Reid's failure was made worse by his success in turning British

---

[13] Laudan, *Science and Hypotheses*, 129.

methodological thought against the method of hypothesis to the wrong-headed inductivism.

In one of the most interesting chapters of his book Laudan refers to Hume's concept of enumerative induction, i.e. induction where a universal empirical generalization is supposed to be warranted by a certain number of positive instances of it, as the *plebian problem of induction*.[14] According to Laudan, it lacks the scientific significance of *aristocratic* induction, which is induction to theories, i.e. induction where a theory or hypothesis is supposed to be warranted by a certain number of *confirming* instances of it (i.e. by its testable deductive consequences).

Confirming instances of a theory need not be its positive instances. For instance, a theory may be confirmed by positive instances of one of its consequents. Also, "it is entirely conceivable that all the possible confirming instances of a theoretical statement could be true and yet the statement could still be false."[15] This is because theoretical statements normally state more than what is said by their confirming instances. Such statements typically talk about unobservable and not directly testable micro-structural processes. For instance, Newton explained Boyle's law by a theory of gases composed of imperceptible particles which repel one another. Likewise, Descartes sought to explain planetary astronomy by assuming that all the planets are carried around the sun by a rotating fluid composed of imperceptible particles, i.e. a vortex.[16]

According to Laudan, this aristocratic concept of inductive evidence existed long before the 1600s, and therefore, there also was a sceptical problem of aristocratic induction long before Hume introduced his plebian problem. Hume's blindness to, and avoidance of, aristocratic induction was, according to Laudan, "probably related to his almost unparallelled ignorance of the science of his time."[17]

Reid, on the other hand, was eminently qualified in the sciences of his time. As Laudan puts it, "Reid was ideally suited to serve as Newton's spokesman. Trained in natural sciences, Reid lectured at Aberdeen and Glasgow in physics and mathematics as well as philosophy in the narrower sense."[18] The question arises, therefore, whether Reid understood the significance of the aristocratic concept of induction.

Although Laudan does not address this question directly, his negative response to it is quite clear. Let me quote him in detail:

> What this [Reid's interpretation of Newton's doctrine of *vera causa*] amounts to is the claim that *unobservable entities*, because we can have no *direct* evidence of their existence, *have no role to play in causal explanations*. In Reid's hands, Newton's first rule of reasoning becomes a vehicle for excluding all theoretical entities from natural philosophy [...] Like Hume before him, Reid will not traffic in theories whose predicates are not drawn directly from sense.

---

[14] *Ibid.*, 72–85.

[15] *Ibid.*, 75.

[16] *Ibid.*, 73.

[17] *Ibid.*, 83.

[18] *Ibid.*, 88.

> Reid uses Newton's *vera causa* doctrine as a tool for discrediting the argument – usually associated with the method of hypothesis – that unobservable causes can be legitimately invoked so long as their existence can be *indirectly* confirmed, i.e. so long as they are capable of explaining what can be observed. Reid's position is that no amount of confirmation or explanatory success is sufficient to establish the existence of an unobserved entity.[19]

So although Reid was more schooled than Hume in natural sciences, he seems to have been as blind as his famous predecessor to the aristocratic concept of induction. Or rather than just blind he seems to have been openly hostile to it. Why was that?

The received view has it that Reid was an avid and somewhat dogmatic supporter of Newton, who had a very simplistic understanding of Newton's *Regula Philosophandi*. He was always eager to point out Newton's attacks against hypotheses but ignored the fact that Newton also extensively utilized the method of hypothesis (as in the explanation of Boyle's law).[20] Together with Hume, he failed to understand the "aristocratic" hypothetico-inductive model of scientific thinking and hence opted for the simple-minded and observation-based enumerative inductivism.

This view of Reid, endorsed by Laudan, as a "plebian" promoter of Newton's inductivism is completely false. It is based not only on a misinterpretation of Reid but also on some fundamentally misconceived ideas of Newton's methodology. It is true that Reid was hostile to aristocratic induction, but not for the reasons Laudan thinks. In particular, he was not motivated by any sort of plebian or enumerative concept of induction. He had plenty of more fundamental reasons to be skeptical of it.

First of all, Reid was motivated by the assumption that nothing could justify hypotheses that were originally of the wrong type. Aristocratic induction does not change the fact that, as Laudan himself puts it, "a theory should not be regarded as confirmed or well established merely because it was sufficient to save the appearances."[21]

Laudan makes a distinction between two different types of hypothesis, *artificial* and *legitimate*; the former referring to hypotheses that merely aim to save the known phenomena, the latter to hypotheses that genuinely aim to go beyond their initial data base for their confirmation.[22] An artificial hypothesis cannot be confirmed by aristocratic induction because the truth of a hypothesis is totally *underdetermined* by the confirming data base. Not only does the same data allow several competing hypotheses, among which there is no way of rationally choosing just the right one, but also such a hypothesis can never be shown to be *false* by any empirical happening whatever. It may provide a satisfying picture of facts already known, but it

---

[19] *Ibid.*, 93.

[20] For Newton's views on hypothesis see e.g. Alan E. Shapiro, "Newton's Optics and Atomism," in I. Bernard Cohen and George E. Smith (eds.), *The Cambridge Companion to Newton* (Cambridge: Cambridge University Press, 2002), 227–55.

[21] Laudan, *Science and Hypotheses*, 126.

[22] *Ibid.*, 127–35.

cannot make predictions, and henceforth, gives no empirical information and does not state any scientific fact.[23]

In other words, aristocratic "induction to theories" is of no use if the underlying hypotheses are artificial, and the problem was that in Reid's time this was the case with all hypotheses proposed by those "votaries of hypotheses" that Reid was criticizing. As Reid himself puts it:

> If the hypothesis hangs well together, is embellished by a lively imagination, and serves to account for common appearances, it is considered by many as having all the qualities that should recommend it to our belief, and all that ought to be required in a philosophical system (IP, I, III, 235a).

It is important to recognize that the early hypotheticalism of LeSage, Hartley, and Lambert was of the artificial kind. As we compare them with the eighteenth-century empiricists, such as Reid and Hume, and evaluate the empiricists' lack of understanding of, or enthusiasm for, the pioneering spirit of the early hypotheticalists, the recognition of these insufficiencies in the views of LeSage, Hartley, and Lambert is critical in order to understand why Reid and the other inductivists were so adamant in their defence of the Newtonian methodology against the way of hypotheses. Laudan quite rightly points out that

> [i]n stressing that a hypothesis must establish its credentials by going beyond its initial data base, Herschel and Whewell defined a promising *via media* between the earlier extremes of Hartley and Le Sage, on the one hand, and Reid and the other inductivists, on the other.[24]

However, this equal attribution of blame to both sides could be considered far too modest a concession to the critical acumen of the eighteenth-century empiricists, especially in reference to Reid's additional observation that many of the theories of the early hypotheticalists do not save even the known phenomena. One such theory was Hartley's theory of the mind, according to which sensations and ideas are caused by "miniature vibrations of the medullary substance of the nerves and brain" called vibratiuncles.[25] According to Reid, such a theory cannot explain the known variety of different sensations – to say nothing of its possessing any predictive power. (IP, II, III, 252a–b)

Most importantly, however, there are serious problems not only with artificial hypotheses but also with "legitimate" hypotheses of aristocratic induction (or con-

---

[23] See e.g. Mary B. Hesse, *Forces and Fields* (London: T. Nelson, 1961), Chapter I. Hesse calls artificial hypotheses "*ad hoc* postulates". She writes: "if an opponent of Newton's gravitational theory had wished to maintain that it is really the intelligent souls of the planets which direct them in their orbits, no future observation of the orbits could show this statement to be false, since it would always be open to anyone who asserted it to say that the planetary souls had willed whatever orbits happened to be observed" (9). Hesse also specifies various forms of artificiality – for instance, she writes: "Descartes's vortex theory is an example of vagueness: almost any physical happening could be described in terms of combinations of vortices in the aether, and Descartes himself showed considerable ingenuity in inventing several such descriptions" (10).

[24] Laudan, *Science and Hypotheses*, 129.

[25] I shall return to some details this theory in Section 7.4. below.

firmation theory). For are these hypotheses, that go beyond their initial data base for their confirmation, mere randomly formed wild guesses, or are they products of a systematic thinking leading to a proper heuristics? This is the 64 million dollar question for the hypotheticalists who have tried to answer it in quite many different ways (Popper's answer being perhaps the most famous negative one). It is quite clear that, according to Reid, such heuristics is needed for a justified use of hypotheses in science. Not just any hypotheses or conjectures will do for Reid. Acceptable hypotheses must be products of a logic, and according to Reid, this logic is inductive. It is here that it may look like the contexts of justification and discovery are getting somewhat mixed up in Reid's philosophy, but the truth of the matter is that the confusion lies rather in the mind of the modern reader than in Reid's thought. Reid's induction is clearly "induction to theories", but it is not so much a method of confirming and justifying those theories on the basis of some "legal" pattern of inference as it is a method of their discovery. As it happens, this method was also well known to many medieval and ancient thinkers.

To use Laudan's own catchphrase, many of the "legitimate" hypotheses of aristocratic induction are, in fact, too "plebian" for Reid. The most important and characteristic feature of Reid's inductivism relates to the fact that neither Reid nor Newton would have considered the hypothetico-inductive confirmation theory genuinely *inductive* at all. Laudan's claim that this concept of induction-*cum*-confirmation "existed long before the 1600s" is, in fact, highly dubious. It did exist, of course, but not as a theory of induction. Confirmation of theories by their testable deductive consequences belonged to the latter part of the classical method of analysis and synthesis. Induction, on the other hand, was part of the method of analysis. It dealt with generalization once hypotheses had been first "analyzed" into their ingredients. Such induction had very little to do with confirmation. It was a method of discovery, not a method of justification.

Reid's most explicit endorsement of this classical view of induction is in a letter to Lord Kames where he writes that "the whole theory of natural philosophy" is made up of *analytical*, *synthetical*, and *practical* parts. The analysis consists of "just and cautious induction" by means of which we "discover" laws of nature. The synthetical part takes for granted the causes discovered in the analysis and uses them to explain the phenomena which result from them. "The practical part consists in applying the laws of nature to produce effects useful in life."[26]

The classical concept of induction as an instrument of discovery was hailed by Newton, as has been shown by various scholars.[27] It should also be noticed that this classical concept of induction entails a concept of hypothesis and that it was this concept of hypothesis which was endorsed by both Newton and Reid as they continued their attacks on those uses of hypotheses (either artificial or legitimate in Laudan's opinion) that did not follow the model set by this concept. So what Laudan

---

[26] See letter to Lord Kames (16th December 1780) in *Philosophical Works*, 57a.

[27] Perhaps most notably by Jaakko Hintikka and Unto Remes in their *The Method of Analysis* (Dordrecht: D. Reidel, 1974), 105–17.

seems to miss is that Reid had good reasons to reject not only "artificial" hypotheses but also many of those hypotheses that Laudan calls "legitimate". For neither Reid nor Newton would have accepted that a hypothesis could be legitimate solely by virtue of having testable deductive consequences. This is the concept of hypothesis mainly endorsed by the modern-day hypothetico-deductivists (or hypothetico-inductivists), but it is not the classical concept endorsed by Newton and Reid. As Jaakko Hintikka and Unto Remes have persuasively argued:

> Newton's method differs from the usual descriptions of the so-called hypothetico-deductive method precisely in the way his strictures against hypotheses bring out. For him, not any hypothesis having testable deductive consequences is acceptable. He does not allow any old hypotheses in our sense of the word, but only those that have been inferred, derived, or, as he sometimes puts it, 'deduced' from phenomena. He is not only concerned to provide a set of rules for confirming hypotheses and theories already known, but also to outline a rational method of finding these hypotheses and theories in the first place.[28]

In other words, many of Laudan's legitimate hypotheses are, in fact, "plebian" in the sense that they are not properly developed or found in the context of the experimental setting in question. Instead they are just any conjectures and guesses which happen to have testable deductive consequences (as such they are better, of course, than the artificial hypotheses of LeSage, Hartley, and Lambert that cannot be tested at all). Reid's (and Newton's) hypotheses are "fine imaginations and most ingenious conjectures." They are well-staged predictions that are intimately linked with the experimental setup at hand. They propose proper questions to nature whose answers may then lend support to their truth. They are anticipated by nature's earlier answers in such a way that in some cases they amount to a type of extrapolation. Reid's hypotheses concerning squinting in the *Inquiry* serve as good examples of such "aristocratic" hypotheses.

These "aristocratic" hypotheses are conceptually related to the constructions and theorems (or "solutions") of the analytical method of ancient geometers. The method has been shown to have been instrumental in the works of Newton.[29] In fact, the simplest form of aristocratic hypothesis is seen in action in Newton's theory of gravity in which the limited generalization of the "centripetal" force (i.e. the force by which satellites – like the moon – are kept in their orbits) is extended to all planets, i.e. also planets without satellites.[30]

In conclusion, from the point of view of the history and development of scientific methodology the boundaries between hypotheticalism and inductivism are not as clear-cut as Laudan seems to think. Hypotheticalists (of any type) have not always been on one side, as it were, and the inductivists (of any type) on the other. On the contrary, the attack by eighteenth-century inductivists (such as Reid) on plebeian hypotheses (either artificial or legitimate in Laudan's terms) was motivated by a

---

[28] *Ibid*., 110.

[29] See Hintikka and Remes, *The Method of Analysis*, 105–17.

[30] See William Harper, "Newton's Argument for Universal Gravitation," in I. Bernard Cohen and George E. Smith (eds.), *The Cambridge Companion to Newton* (Cambridge: Cambridge University Press, 2002), 185–87.

methodology that placed a high importance on well-formed and experimentally supported hypotheses in scientific thinking. The inability of various scholars to see this has been caused by their insufficient understanding of the role of induction in science as an instrument of discovery along the lines of the classical method of analysis.

7.4 Reid's Inductivism and His Philosophy of the Mind

It is the aforementioned inductivist attack on plebeian hypotheses – not on aristocratic ones – that also characterizes Reid's philosophy of the mind. Reid employs a kind of derivative of Ockham's razor. He begins with the question of how the operations of the mind are explained. To answer this question, to find out the explaining causes of our mental phenomena, Reid thinks we must follow two conditions taught us by Newton: first, these causes ought to have real existence, and secondly, they ought to suffice as explanations, i.e. be able to cause what they are supposed to cause.[31]

The problem with plebeian hypotheses is that they tend to lead to an ontological slum of explaining causes. Because causes ought to have real existence, the hypotheticalists assume that the causes they have conjectured do exist. But this leads to a reification of causes without any limit. For there is no way of saying of competing plebeian-*cum*-artificial hypotheses which particular one really presents true existing causes. Reid thinks that one should try to avoid this jungle of existing causes and that such causes should not be multiplied beyond necessity.[32]

Reid tries to show the hypotheticalists that their methodology is arbitrary and that they should – if they want to hold to their hypotheticalism – develop such criteria of discovery that makes a rational evaluation of competing hypotheses possible. At the same time, Reid is, in a sense, promoting a new type of metaphysics of the mind – a metaphysics that is anti-reductionist in its methodological essence. He is an early opponent of reductive system-building in philosophy. His attack on plebeian hypotheses reflects this opposition. As indicated above, the differences between plebeian and aristocratic hypotheses are plenty. One consists in the fact that in going beyond their initial data base for their confirmation, aristocratic hypotheses go *to a data base* for their confirmation (i.e. they are testable). Of the plebian hypotheses only the legitimate ones do this. Artificial hypotheses cannot be properly tested at all. They must be judged on the basis of their initial data base – which, as has been shown, is an insufficient ground of evaluation – and on the basis of such non-data as the grandiosity of the system they establish.

Another difference lies in the fact that for aristocratic hypotheses it does not suffice that they are hypotheses that go to a data base for their confirmation (i.e. have

---

[31] "It is a dictate of common sense, that the causes we assign of appearances ought to be real, and not fictions of human imagination. It is likewise self-evident, that such causes ought to be adequate to the effects that are conceived to be produced by them" (IP, II, VI, 261a).

[32] Cf. e.g. Reid's discussion of gravity in IP, II, VI, 261. Reid uses Newton's theory of gravity as a model of philosophizing for his philosophy of the mind.

testable deductive consequences). They also need to be "deduced" from phenomena (i.e. be derived from the experimental situation at hand). Plebeian hypotheses are not so derived and lead therefore to inflated metaphysics. As Reid puts it:

> Philosophy has been, in all ages, adulterated by [...] systems built partly on facts, and much upon conjecture (IP, II, III, 249b).

Notice that we are referring to Reid's *methodological* anti-reductionism. In other words, he rejects reductionist methods that lead to inflated metaphysics. These methods aim at reducing mental phenomena to some major hypothetical doctrine from which such phenomena are supposedly deduced. These methods are reductionist, of course, but their outcome is not characterized by simplicity, as one might expect. On the contrary, the outcome is often plagued by the multitude of the unexplained, i.e. of metaphysical entities and principles that do not derive from any experimentation, resist natural explanation, and are only supported by the coherence of the system they belong to. Reid wants the opposite. He wants a non-reductionist methodology that produces metaphysical simplicity as a result.[33] To achieve this he adopts an inductivist methodology that consists of accurate experiments and patient observation ascending to general facts and laws of nature. Such non-reductionist methodology minimizes inflationary metaphysics. As Reid would say, it prevents castle-building in the air.[34]

Reid gives plenty of examples of such castle-building in his *Essays on the Intellectual Powers of Man*, in particular, in the chapter called "Hypotheses concerning the nerves and brain". He begins by noting that anatomists inform us that nerves have a fibrous texture, and may be divided and subdivided, till its fibres escape our senses. And he continues:

> [...] as we know very little about the texture of the nerves, there is great room left for those who choose to indulge themselves in conjecture (IP, II, III, 248b).

He then goes on describing various hypothetical theories of how the nerves perform their functions. For instance, the ancients thought that the nerves performed their functions by animal spirits, a kind of vapour that filled the fine tubes they thought the nervous fibres consisted of. This theory was, according to Reid, further developed by Descartes who described the going and returning of these spirits in the

---

[33] Notice that this "metaphysical simplicity" should not be confused with simplicity as a virtue of scientific explanation which Reid considers a major source of error in science. Cf. IP, VI, VIII, 470b–472a.

[34] On the non-reductionist tendency in Reid's philosophy see, for instance, what he says about "evidence of sense and of belief": "The common occasions of life lead us to distinguish evidence into different kinds, [...] such as the evidence of sense, the evidence of memory, the evidence of consciousness, the evidence of testimony, the evidence of axioms, the evidence of reasoning [...] Philosophers have endeavoured, by analysing the different sorts of evidence, to find out some common nature wherein they all agree, and thereby to reduce them all to one [...] I confess that, although I have, as I think, a distinct notion of the different kinds of evidence above-mentioned, and, perhaps, of some others, which it is unnecessary here to enumerate, yet I am not able to find any common nature to which they may all be reduced" (IP, II, XX, 328a–b).

nerves and how this affected everything from muscular motion to mental phenomena. Reid says:

> All this he has described as distinctly as if he had been an eye-witness of all those operations. But it happens that the tubular structure of the nerves was never perceived by the human eye [. . .] and all that has been said about animal spirits, through more than fifteen centuries, is mere conjecture (IP, II, III, 248b).

Reid then comments on the book called *Nova Visionis Theoria* by Dr Briggs, who had been Newton's master in anatomy. Briggs developed a theory according to which the nerves are very thin solid filaments that vibrate like musical cords, thereby performing their functions. Reid notices that this theory was never much followed because the nerves, having a very small degree of consistence and in contact with moist substances, are very unfit for this type of performance (IP, II, III, 248b–249a).

Reid also makes some interesting comments on Newton's views of this subject. He first of all notes that Newton has put his conjectures in the form of queries, so that

> [. . .] they might not be received as truths but be inquired into, and determined according to the evidence to be found for or against them (IP, II, III, 249a).

He then introduces Newton's ether theory according to which there may be an elastic medium, immensely rarer than air, which pervades all bodies and which is the cause of gravitation, refraction and reflection of the rays of light, the transmission of heat, and ultimately also the impressions made on the nerves and brain in perception. The question is whether, for instance, vision is

> [. . .] effected chiefly by the vibrations of this medium, excited in the bottom of the eye by the rays of light, and propagated along the solid, pellucid, and uniform capillaments of the optic nerve (IP, II, III, 248a).

It is via these comments on Newton's queries that Reid is then led to discuss David Hartley, who, according to Reid, made Newton's mere queries into doctrines from which

> [. . .] he has deduced, in a mathematical form, a very ample system concerning the faculties of the mind (IP, II, III, 248a).

As said, Hartley's theory of the mind consists of a doctrine of physical vibrations. He claims that mental and neural events operate on parallel tracks in which neural events cause mental events. On the mental side, we have sensations and ideas. Sensations are caused by vibrations of a fine granular substance within the nerve-tubes of peripheral sensory nerves. Ideas, on the other hand, are caused by miniature vibrations in the brain called vibratiuncles.[35] According to Hartley, all the operations of the mind can be reduced to sensations and ideas that supervene on physical vibrations in the nerves.

According to Reid, Hartley's theory is pure speculation and violates two conditions of scientific explanation taught us by Newton. First, Hartley produces insufficient evidence for the existence of vibratory motions in the nerves and brain (i.e. the

---

[35] Cf. e.g. Thomas H. Leahey's short entry on Hartley in R. Audi (ed.), *The Cambridge Dictionary of Philosophy* (Cambridge: Cambridge University Press, 1995).

hypothesis is not experimentally supported; it is neither "deduced" from phenomena nor derived from any experimental setting), and secondly, these vibratory motions are insufficient to account for all our sensations and ideas, which again are supposed to account for all the operations of the mind.[36] Reid asks how we can expect any proof of the connection between vibrations and thought when the existence of such vibrations was never proved. The proof of their connection cannot be any stronger than the proof of their existence. As to the existence of our thoughts, we have the evidence of our consciousness; but as to the existence of vibrations in the nerves and brain, no proof has yet been brought. And even if such a proof were found, vibrations could still not explain all the operations of the mind. In fact, vibrations cannot even explain the sensations we have by means of our external senses – not to say anything about sensations that accompany our passions and affections. Reid says:

> We have five senses, whose sensations differ totally in kind. By each of these, excepting perhaps that of hearing, we have a variety of sensations, which differ specifically, and not in degree only [. . .] How shall we find varieties in vibrations corresponding to all this variety of sensations which we have by our five senses only? (IP, II, III, 252a–b)

And he continues:

> [. . .] when we are told of vibrations in a substance which no man could ever prove to have vibrations, or to be capable of them; when such imaginary vibrations are brought to account for all our sensations, though we can perceive no correspondence in their variety of kind and degree to the variety of sensations – the connections described in such a system are the creatures of human imagination (IP, II, III, 253a).

Notice, by the way, that all these hypothetical accounts of the mind seem to have in common the mechanistic idea of "cognitive contact". This idea was particularly popular among the philosophers of the early modern era. It makes the mind and its relation to the external world conform, in one way or another, to the famous principle of "no action at a distance". Applied to the science of the mind this principle becomes the principle of "no cognition or cognitive action at a distance". The problem with this principle is that it makes the mind a place in which various mental operations are defined in terms of their four-dimensional spatial relations to each other. Ultimately, the mind itself is so defined in its relation to the external world. Animal spirits, etherial media, and vibrating filaments literally aim at filling in the spatial gaps in the causal chain of cognition.

A clear strength of Reid's inductivism is that it keeps away from such hypostasis of spatial elements in cognitive activity. First of all, Reid has no problem with the idea of action at a distance, physical or cognitive. This follows from his allegiance to Newton's way of doing natural philosophy. According to Reid, "immediacy" and "directness", which both characterize the activities of the mind, do not mean "presence". The distinction between things in the mind and things external is not meant to signify the place of things we speak of, but their subject.[37] The mind is essentially the subject of its operations. It is also from its very nature a living and active being.

---

[36] IP, II, III, 250a–253a.

[37] IP, I, I, 221b.

The philosophers' task is to specify the powers of the mind, its capacities, habits, and dispositions, to examine them in detail, and to draw general conclusions from the facts that such examinations reveal. All this should be done without recourse to metaphysical speculation on the hidden causes of such operations, causes that might be needed for a mechanistic account of mental activity. Of such causes we know nothing, and scientific research is best advanced by our admitting our ignorance of such causes and by our giving up any such principle – such as the principle of cognitive contact – that makes us speculate on the issue. In the final paragraph of the chapter *Hypotheses concerning the nerves and brain* Reid argues:

> The rays of light make an impression upon the optic nerves; but they make none upon the auditory or olfactory. The vibrations of the air make an impression upon the auditory nerves; but none upon the optic or the olfactory. The effluvia of bodies make an impression upon the olfactory nerves; but make none upon the optic and auditory. No man has been able to give a shadow of reason for this. While this is the case, is it not better to confess our ignorance of the nature of those impressions made upon the nerves and brain in perception, than to flatter our pride with the conceit of knowledge which we have not, and to adulterate philosophy with the spurious brood of hypotheses (IP, II, III, 253a).

In conclusion, Reid's inductivist method served an important scientific purpose in the eighteenth-century debate on the character and content of human understanding. Reid's methodological ideas are closely intertwined with his attack on the theories of the mind that, on the one hand, still dealt with spirits and souls and, on the other, were flooded by spurious mechanistic hypotheses. Commentaries that focus exclusively on Reid's methodological ideas tend to miss the essential point that as much as on the methods of science, Reid was still working on the methods of metaphysics. Or better still, he was conducting a demarcation dispute between the methods of science and metaphysics,[38] searching for a philosophy that would transform parts of traditional metaphysics into a modern science, i.e. the doctrine of souls and spirits into a scientific mental philosophy.

7.5 Some Weaknesses of Reid's Inductivism

Let me now say something about the possible weaknesses of inductivism in Reid's theory of the mind. One of them is Reid's tendency to exaggerate the reification of hypothetical entities by those philosophers that do not abide by his inductive rules of reasoning. The most famous example of such exaggeration relates to Reid's attack on the theory of images and ideas. It has been said that no philosopher has ever held that version of the theory of ideas that Reid criticizes. Although this is not a totally fair description of Reid's position, there is a grain of truth in it.

In that version of the theory of ideas that Reid criticizes, the word "idea" is given a philosophical meaning according to which it does not merely signify (as it does in its popular meaning) that act of mind which we call thought or concept formation,

---

[38] See, in particular, Robert Callergård's *An Essay on Thomas Reid's Philosophy of Science* on the demarcation issue.

but some independently existing, non-material object of thought. In this interpretation ideas are not just mere contents intrinsic to cognitive acts but immediate objects of cognition that have objective existence in the mind. They are "things" or "entities". They are substance-like proxy objects. According to Reid, this hypostasis of ideas as dummy objects has misled many philosophers; it is, for instance, the root cause of Berkeley's immaterialism and Hume's scepticism.[39]

The question is, therefore: did any of the philosophers Reid talks about ever hold this theory of ideas as "entities"? The answer is that some did at some point, but that no single position can be associated with the phrase "theory of ideas" at any point of time. It has been said that Berkeley's theory fostered the mistaken idea that all philosophers shared a single theory of ideas and that Hume's theory posited ideas that can possibly exist independently apart from any mind. In his 1979 paper, "Locke's Mental Atomism and the Classification of Ideas: I," M.A. Stewart, on the other hand, talks famously about "those dummy objects of that dummy Locke beloved of the critics" – an interpretation of Locke's position that has been seconded for instance by Reginald Jackson, R.I. Aaron, H.A. Prichard, and D.J. O'Connor. A.D. Woozley and John Yolton, however, hold a different view of Locke. They think that Locke's ideas are perceptions. Locke talks about "having" ideas like having thoughts, rather than "seeing" them, i.e. viewing them like objects. Most importantly, however, Ian Tipton shows in his 1999 paper that neither of these two interpretative approaches has any well-defined borders. It is not only that Locke's position is ambiguous, so are the primary clarifications of it by the commentators.[40]

The truth seems to be that philosophers have supported various interpretations of the theory of ideas and have found it difficult to decide which interpretation they really prefer. It seems that the theory of ideas was seen as such a powerful tool of explanation that although many of its supporters were perfectly aware of its ambiguities, they were still ready to accept it and to downplay its problems. So, for instance, Locke complains in a letter to John Norris that "if you once mention ideas you must be presently called to an account *what kind of things you make these same ideas to be*."[41]

Notwithstanding these reservations, however, it is safe to say that Reid's estimation of the popularity of the "proxy" theory of ideas is exaggerated. As one studies his interpretations of other philosophers' theories, one cannot but think that the popularity of the "proxy" theory is only a hypothesis that Reid is far too fond of and that he should have directed the following words that appear in *Essays on the Intellectual Powers of Man* to himself:

> When a man has, with labour and ingenuity, wrought up an hypothesis into a system, he contracts a fondness for it, which is apt to warp the best judgment (IP, II, III, 250b).

---

[39] Cf. IP, II, XIV

[40] This quick review of Locke commentary is based on Ian Tipton, " 'Ideas' and 'Objects': Locke on Perceiving 'Things'," in M. Atherton (ed.), *The Empiricists: Critical Essays on Locke, Berkeley, and Hume* (Oxford: Oxford University Press, 1999) where the exact references are to be found.

[41] Quoted in Tipton, " 'Ideas' and 'Objects': Locke on Perceiving 'Things' ".

But, at the same time, it needs to be observed that the real target of Reid's criticism is, perhaps, the very ambiguity in the views of the "idealists". His description of Arnauld's theory, for instance, suggests that his criticism is not only directed at those who, he thinks, prefer a "proxy" theory of ideas to some other interpretation, but also at those who, like Arnauld, keep on entertaining this theory instead of clearly rejecting it as he has done.[42] In other words, Reid thinks that philosophers must either give an exact account of what kind of things they make ideas be and give evidence for their existence according to the status of their being, or, in case they cannot do this, give up the theory of ideas unequivocally.

Another possible weakness of Reid's inductivist theory of the mind is that when it comes to enumerating those particular facts that form the basis of inductive generalization about the laws of the mind, Reid seems to think that "the unshakable beliefs of all mankind" are such facts. We can try to understand this common sense philosophy in at least two different ways; i.e. more or less kindly. First, it can be judged against the metaphysical background of Reid's views. Common sense is, after all, an efficient tool against metaphysics that is plagued by plebeian hypotheses. Reid could be understood as saying that inductive science of the mind must start from the unshakable beliefs of all mankind once all plebeian hypotheses are rejected. We must start, as it were, from a clear table, void of all metaphysical obscurity. It is then up to empirical investigation and a careful examination of all the facts to contradict these beliefs if needed. Common sense is the starting-point of all science, but not something that must be obeyed against all evidence.

Another favourable interpretation of these beliefs or "principles" of common sense – and which in a way complements the previous one – connects them with medieval thinkers' "general propositions". Like Duns Scotus, Reid was an onto-logical nominalist.[43] He needed nontrivial initial premises to back up his particular observations and facts in order to reach universal or general conclusions about nature. One such premise shared by Duns Scotus and Reid was the principle of the regularity of nature. Such an interpretation of Reid's first principles goes well with his allegiance to the interrogative model of inquiry as shown by his use of "putting questions to nature" terminology. In addition, it would explain his attack on Hume's sceptical brand of nominalism where the lack of such initial propositions seems to result in a collapse of the classical (i.e. Aristotle's) model of inquiry and in a transformation of epistemology into an exclusive debate about verification.[44]

---

[42] Cf. the entry on "Idea" by M.B. Bolton in J. Kim and E. Sosa (eds.), *A Companion to Metaphysics* (Oxford: Oxford University Press, 1995).

[43] The exact brand of Reid's nominalism is of no importance here.

[44] The "initial premises" view of the medieval thinkers and its role and position in the history of epistemology and methodology is described and investigated by Jaakko Hintikka in his "Concepts of Scientific Method from Aristotle to Newton," in Monika Asztalos, John E. Murdoch, and Ilkka Niiniluoto (eds.), *Knowledge and the Sciences in Medieval Philosophy*, Acta Philosophica Fennica, vol. 48, Proceedings of the Eighth International Congress of Medieval Philosophy, Helsinki 1990, 72–84. Hintikka points out that because of these initial premises "nominalism cannot be construed as a skeptical philosophy" (*ibid*., 77), as is often done by its critics.

But we can also give a less generous interpretation of Reid's common sense philosophy. We can say that his appeal to the beliefs of all mankind, to "universal consent in things not deep nor intricate," reveals his inability to understand the significance of bold and extravagant theorizing in scientific thinking. His repeated appeals to ignorance and its worthiness in the face of valid problems of science – his preference to surrender rather than to conjecture – reveal his debilitating religious prejudices. He appeals too easily to God and the unreachable mysteries of the Supreme Being to be taken seriously as a philosopher. Of everything that goes beyond common sense Reid is ready to say:

> [. . .] our Maker has limited and circumscribed our powers of perception, by such laws of Nature as to his wisdom seemed meet, and such as suited our rank in his creation (IP, II, IV, 257b).

If, however, the virtues of a theory are based on its ability to dispense with objections, then the aforementioned favourable interpretation, according to which "the beliefs of all mankind" are initial premises of erotetic inquiry, is supported by the following consideration.

It has been objected that Reid's appeal to common sense actually threatens his inductivism and that the relationship between Reid's inductivist methodology and his common sense philosophy is, therefore, permanently tense and problematic. For rather than expressions of particular facts of human psychology the "unshakeable beliefs of all mankind" are general principles, the relation of which to Reid's inductive method is fundamentally different from any relation that particular facts can hold to it. If these principles are general, do they not undermine Reid's inductivism, according to which such principles should be attained inductively and not by assumption? What exactly is the role of these principles *vis-à-vis* Reid's inductivism?

Let us first look at Reid's common sense beliefs in more detail. They include such beliefs as

i. "the existence of everything of which I am conscious"
ii. "those things did really happen which I distinctly remember"
iii. "the thoughts of which I am conscious are the thoughts of a being which I call myself, my mind, my person"
iv. "the natural faculties, by which we distinguish truth from error, are not fallacious"
v. "in the phenomena of nature, what is to be, will probably be like what has been in the similar circumstances"[45]

In the end, Reid argues that we ought to take for granted all "things wherein we find an universal agreement, among the learned and the unlearned, in the different nations and ages of the world." And to those who argue that "it is impossible to collect the opinions of all men upon any point whatsoever" he answers that "there

---

[45] Cf. IP, V.

are many cases wherein it is otherwise," i.e. the universality of certain opinions is "sufficiently evident" to be taken for granted.[46]

What Reid could be saying here is that certain individual beliefs that seem to be held by "the learned and the unlearned" can be generalized into beliefs held by all men in different nations and ages of the world. The universality of such beliefs is "sufficiently evident", and this evidence is inductive (i.e. the proposition that beliefs are held by a certain number of "the learned and the unlearned" is generalized into the proposition that they are held by all men). According to this interpretation, Reid's common sense principles could be something like inductively discovered general propositions, the universality of which is also inductively justified.[47]

But from this inductively grounded universality of beliefs Reid also seems to conclude that the instances of such universal generalizations are *self-evident*[48] or that every individual common sense opinion is an instantiation of a *self-evident* principle. Where does this self-evidence come from? What does it mean? How do we get from inductively grounded universality of beliefs to their self-evidence, or better still, to the self-evidence of their propositional content?

It can be argued that Reid's notion of self-evidence is more psychological and pragmatic than epistemic or analytic.[49] Common sense principles are self-evident because they are psychologically compelling (i.e. they are natural and instinctive belief generators), and because they are, as Reid says, "of infinite advantage upon the whole" (**I**, VI, XXIV, 199b). Their self-evidence means that they are an inescapable part of human nature. As such, this self-evidence has nothing to do with the inductive grounding of their universality.[50]

Although Reid does not explicitly say that self-evidence is a psychological notion, various examples of the formation of individual common sense beliefs and of the role of the notion of self-evidence in this process indicate that the meaning of self-evidence lies in the intrinsic nature of instinctive and innate belief-forming processes rather than in intuitive knowledge – in the sense of intellectual or rational

---

[46] Cf. IP, I, II.

[47] Remember also what Reid says about particular beliefs *vis-à-vis* their generalizations: "It is another property of [...] many first principles, that they force assent in particular instances, more powerfully than when they are turned into a general proposition" (IP, VI, V, 448a).

[48] Cf. IP, VI, IV, 434b.

[49] In my 1993 study of Reid I argued explicitly that Reid's notion of self-evidence is best understood as some kind of psychological "ground of irresistible belief"; see *Pioneering Externalism. Thomas Reid and Hume's Problem* (Reports from the Department of Philosophy, University of Helsinki. No 1, 1993), 162–72. Cf. also C. Engel-Tiercelin, "Reid and Peirce on Belief," in M. Dalgarno and E. Matthews (eds.), *The Philosophy of Thomas Reid*, (Dordrecht: Kluwer Academic Publishers, 1989), 205–24. Engel-Tiercelin focuses on Reid's theory of belief-formation rather than on his theory of evidence, but these two theories are closely intertwined. Engel-Tiercelin describes Reid's theory of belief as follows; "it [i.e. belief] has nothing to owe to reason or experience, it is a principle of our constitution, connected with instinct" (*Ibid.*, 206).

[50] The curiosity of Reid's notion of self-evidence is perhaps best epitomised by his somewhat puzzling concept of "self-evident probability". In general, for Reid, self-evidence does not entail either absolute certainty or necessity (IP, VI, IV, 435 a–b).

insight into their propositional content. The lack of any explicit definition on Reid's part could be explained by the insufficient differentiation between the psychological and the epistemic in his philosophy. This insufficiency is demonstrated e.g. by the following quotations:

> But the power of judging in self-evident propositions, which are clearly understood, may be compared to the power of swallowing our food. It is purely natural, and therefore common to the learned and the unlearned, to the trained and the untrained (IP, VI, IV, 434b).
>
> [...] the belief we have, that the persons about us are living and intelligent beings, is a belief for which, perhaps, we can give some reason, when we are able to reason; but we had this belief before we could reason, and before we could learn it by instruction. It seems, therefore, to be an immediate effect of our constitution (IP, VI, V, 441a).
>
> [...] when an opinion is so necessary in the conduct of life, that, without the belief of it, a man must be led into a thousand absurdities in practice, such an opinion, when we can give no other reason for it, may safely be taken for a first principle (IP, VI, V, 441a).
>
> [...] the testimony of memory, like that of consciousness, is immediate; it claims our assent upon its own authority (IP, VI, V, 444b).
>
> If any man asks a proof of [the principle of mind, self, and person], I confess I can give none; there is an evidence in the proposition itself which I am unable to resist (IP, VI, V, 443b).

But things might get more difficult due to the fact that one of the self-evident principles of common sense that Reid discusses is the principle of induction itself. Reid talks about the principles of *veracity* and *credulity* that apply to human testimony as well as to the testimony of nature. The principle of veracity makes us tell the truth (human testimony), and it makes nature follow a regular course (the testimony of nature). The principle of credulity makes us believe what others tell us (human testimony), and it also makes us rely upon the stability of the course of nature or "rely upon the continuance of the connections which experience hath discovered" (the testimony of nature). It is this last application of the principle of credulity – credulity as prescience that things that have been found "conjoined in time past, will be conjoined in time to come" – that Reid specifically calls "the inductive principle".

Again, this inductive principle could be thought of as a generalization from the beliefs and behaviour of "the learned and the unlearned". Its universality and usefulness would be thereby established, and it would be credited with the status of the psychological self-evidence of common sense principles. Thus, no contradiction would ensue between the inductive principle and inductivist methodology. The self-evidence of the principle would be psychological in character and, therefore, not epistemically anti-inductivist. Only the universality of the principle would be provided with an epistemic foundation.

The interesting point is that these suggestions support the aforementioned "initial premises" view of the first principles. For this view says that induction is first and foremost a method of discovery. The problem of the epistemic justification, or the lack of it, of the first principles cannot undermine the method of induction because inductive procedures are not meant to deal with justification in the first place. Justification (verification) in the post-Humean sense of the problem of induction plays no role in Reid's inductivism. Initial premises are basic theoretical assumptions that are part and parcel of the inductive method of discovery. The perceived contradiction

or "tension" between these two constituents of the method of analysis (i.e. initial premisses and induction) is an illusion created by post-Reidian developments in epistemology (the details of which are beyond the scope of this chapter).

Let us recapitulate. On the face of it, Reid's common sense philosophy seems to contradict his methodological inductivism. The former looks like a rationalistic enterprise that compromises the empirical principles of the latter. The former seems to rely on the light of reason and intuition – or, to borrow Reid's own phrase, on "human imagination" – the latter on the observation of particular facts and their generalization. In close study, however, these dichotomies turn out to be either misleading or false. It is only the universality of the common sense principles that is inductively justified. Their self-evidence, on the other hand, means that we have a natural and irresistible impulse to believe in these principles and their instantiations. Because Reid's notion of self-evidence is not an epistemic notion in the contemporary sense of "epistemic" and because his appeal to the self-evident principles of common sense does not invoke a theory of justification in the contemporary sense of the term, there is no contradiction between Reid's common sense philosophy and his methodological inductivism. On the contrary, the function of the self-evident first principles is equivalent to the function of the initial premisses of medieval nominalism. Their task is not related to the verification of inductive procedures in the sense of contemporary theories of epistemic justification but to the functioning of these procedures in a heuristic context.

Chapter 8
Kant on Consciousness

Camilla Serck-Hanssen

Abstract The chapter examines Kant's conceptions of consciousness and their relation to his views on psychology as a science. Kant does not develop a philosophy of mind as such, but through his reinterpretation of metaphysics he develops different notions of consciousness. The most central and specifically Kantian concept of consciousness is that of apperception. It is argued that 'apperception' is not to be understood as self-consciousness or self-awareness. Rather, apperception is a capacity to be aware of one's spontaneous activities, and it can be further analyzed as the ability to respond to rules and norms. Thus understood, 'apperception' plays a central role not only in Kant's theoretical philosophy but also in his moral and aesthetic theory. 'Inner sense' is another central concept for Kant. In the first *Critique* and later works, Kant distinguishes between apperception and inner sense: inner sense is the consciousness of what takes place within the mind as opposed to apperception, which is the consciousness of one's activities. These two concepts of consciousness, 'inner sense' and 'apperception', generate two very different questions about the relation between consciousness and nature. On the one hand, there is the question of how inner or mental nature is related to physical nature; on the other hand, there is the question of how spontaneity is related to the whole of nature, inner nature as well as outer. As we shall see, Kant's answer to the first question is closely related to his pessimism regarding the status of psychology as a science.

Kant never wrote a philosophy of consciousness or a philosophy of mind as such, but this does not mean that he has nothing to say on the topic.[1] In his attempt to establish a critical metaphysics, he develops two different notions of consciousness:

C. Serck-Hanssen
Department of Philosophy, University of Oslo

[1] As we shall see, his way of dealing with the topic is nevertheless quite different from contemporary discussions of mind and consciousness. Cf., e.g., Jaegwon Kim, *Philosophy of Mind*, Dimension of Philosophy Series (Boulder: Westview, 1996).

'apperception' and 'inner sense'. The most central and specifically Kantian concept of consciousness is that of apperception. The first part of this chapter is devoted to this concept. The other concept of consciousness, 'inner sense', will be analyzed in part two.

From Kant's understanding of inner sense and apperception, two quite distinct versions of the mind–body problem arise; these are addressed in the third and last part of this chapter. One version of the problem is how the mental and the physical can affect each other. Kant's views on inner sense both commit him to the existence of psychophysical laws and, somewhat paradoxically, entail that psychology can never become a science proper. The other version of the mind–body problem is how spontaneity, which falls outside the domain of causal laws and nature altogether, can bring about an event in nature (be it mental or physical). In Kant's view we can never know or understand *how* this is possible, but, as we shall see, he nevertheless believes that there are ways of showing *that* it is so.

8.1 Apperception

It is often assumed that "apperception" simply means self-consciousness, but this assumption immediately leads to a problem.[2] Since at least according to one sense of "apperception", this kind of consciousness is supposed to have a transcendental status, i.e., to be an a priori condition of experience, its object cannot be the empirical self. However, since Kant emphatically denies that we have any access to noumena, the noumenal self appears to be ruled out as well. In other words, it seems that the self of apperception cannot be either of the two selves offered by the Kantian framework.[3] What then is the alternative? I believe our best strategy is to begin by looking at the unpublished fragments in which the first systematic occurrences of the word apperception are found.

---

[2] This problem has been acknowledged by several interpreters. Cf., e.g., Patricia Kitcher, "Kant's Real Self," in Allen W. Wood (ed.), *Self and Nature in Kant's Philosophy* (Ithaca and London: Cornell University Press, 1984), 113–47.

[3] Kitcher argues (op.cit.) that the self is nothing over and above a connection of representations. It is, however, hard to see what distinguishes this self from what Kant calls the empirical self. In *Kant's Theory of Mind* (Oxford: Clarendon Press, 1982) Karl Ameriks takes quite a different line. He argues that Kant's contentions in the Paralogisms are not intended to refute claims about a noumenal self, such as those of the substantiality and immateriality of the soul. The targets of Kant's arguments are, rather, the specific and, in his view, illicit ways that rational psychology attempts to establish its conclusions. Since Kant's own claims about the self of apperception are not derived from these faulty inferences, it follows from Ameriks' arguments that there is no reason why this self is not conceived by Kant as the noumenal one. I find it hard to see how this position can be defended. After all, it is not only in the Paralogisms that Kant rejects that the "I" implies that we have access to noumenal selves. Most noteworthy in this connection is § 25 of the Deduction where Kant says (B 157) that in the original synthetic unity of apperception I am neither aware of myself as I appear nor as I am in myself (i.e., as a noumenon).

A common assumption is that before the beginning of the 1770s Kant went through a pre-critical phase during which he was a close follower of Wolff and Leibniz. The accuracy of this view need not concern us here,[4] but it is interesting to note that Kant did not use the Leibnizian term apperception until the so-called silent decade, the years after the *Inaugural Dissertation* (1770) during which Kant published nothing and spent all his free time working on the *Critique of Pure Reason*.[5] More precisely, "apperception" first appears in a collection of fragments from 1774–1775 known as the *Duisburg'sche Nachlass*.[6] Here, apperception is described somewhat confusingly in three different ways: (i) "the perception of oneself as a thinking being in general" (R 4674, Ak 17:647), (ii) "the consciousness of thinking, i.e., of the ways representations become posited in the mind" (R 4674, Ak 17:647) and (iii) "an intuition of the self" (R 4675, Ak 17:651). However, since the object of which one is conscious according to (i) is the thinking being *in general*, (i) does not seem substantially different from (ii). After all, since no particular thinker is hereby individuated, to be conscious of oneself as a thinking being *überhaupt* seems to be nothing over and above being conscious of thinking. Hence, the significant difference between these ways of describing apperception is really between (i) and (ii), on the one hand, and (iii) on the other. According to (i) and (ii), apperception is consciousness or awareness of thinking, an activity which Kant further analyzes as the *ways* in which representations become posited in the mind. Since Kant defines the ways of positing in terms of certain functions or rules,[7] it seems that one could regard apperception as consciousness of the rules that define the activity of such positing or thinking.[8] According to (iii) apperception is an intuition of a self; as such,

---

[4] For an interesting criticism of the standard view see F. Beiser "Kant's Intellectual Development: 1746–1781," in Paul Guyer (ed.), *The Cambridge Companion to Kant* (Cambridge: Cambridge University Press 1992), 26–61. See also C. Serck-Hanssen "Kant's Critical Debut," in Jeff Malpas (ed.), *From Kant to Davidson: Philosophy and the Idea of the Transcendental* (London: Routledge, 2003), 7–22.

[5] References to *Kritik der reinen Vernunft* (KrV) are to the standard A and B pagination of the first and second editions. All other references to Kant are to the volume and page of the Academy Edition (Ak), *Kants gesammelte Schriften* (Berlin: Königlichen Preußischen Akademie der Wissenschaften, 1910). Quotations from *KrV* are taken from N. Kemp Smith's translation (London: MacMillan, 1990), quotations from *Kritik der praktischen Vernunft* (KpV) are taken from M.Gregor's translation (Cambridge: Cambridge University Press, 2002), quotations from *Kritik der Urteilskraft* (KU) are taken from W. Pluhar's translation (Indianapolis: Hackett, 1991), quotations from *Anthropologie in pragmatischer Hinsicht* (Anthro) are taken from M. Gregor's translation (Den Hague: Martinus Nijhoff, 1974). The translations of the unpublished material (*Reflexionen*) are my own.

[6] "Apperception" is also used twice in R 4652, but Addicke's dating of this fragment as being written earlier than the *Duisburg'sche Nachlass* is not certain.

[7] At different places in the *Duisburg'sche Nachlass* and other fragments from the same period these ways are called: "the functions of apperception" (R 4674, Ak 17:646), "the acts of thinking" (R 4674, Ak 17:647, R 4674, Ak 17:649), "the rules of thought" (R 4674, Ak 17:647) and "the titles of thought" (R 4672, Ak 17:635).

[8] Strictly speaking, Kant distinguishes between the rules of thought and the rules of positing, but since we now only deal with apperception from the subjective perspective of consciousness we can overlook that distinction.

it is most likely what Kant later describes as "consciousness of oneself according to the determination of one's states in inner perception [. . .] which is usually called inner sense or empirical apperception" (*KrV*, A 107).[9]

Kant's sudden interest in apperception is undoubtedly intimately connected to his project of showing that certain metaphysical concepts, later identified as categories such as 'cause' and 'substance', can have a real and not only a logical use.[10] If concepts have a real use they not only express the ways in which thinking occurs (this is their logical use), they also contribute to genuine knowledge of objects. In the *Duisburg'sche Nachlass*, we find an argument that appears to be Kant's first attempt to establish the real use of the metaphysical concepts (or categories) through a transcendental deduction that draws on the concept of apperception:[11] "The condition of all apperception is the unity of the thinking subject. From this all connection flows (of the manifold) in accordance with a rule and in a whole [. . .]" (R 4675 Ak 17:651). Just how this argument is supposed to show that there is a real use of the metaphysical concepts, need not concern us in our present context. It is, however, of interest to notice that the crucial interpretational question is how to understand "unity of the thinking subject".[12] As we shall soon see, the concept of such a unity also plays a pivotal role in the transcendental deduction of the first *Critique*.

Most of Kant's published views on apperception are found in the transcendental deduction of the first *Critique*. In the A-edition, Kant presents pure, original, or transcendental apperception as a transcendental kind of consciousness that serves as an a priori condition of all empirical consciousness, i.e., the awareness of objects in space and time. This transcendental consciousness is partly described as self-consciousness (A 107, 117n) or consciousness of one's own identity (A 108), and partly as the *ability* to form the representation 'I' (A 117n), an ability that in its turn is characterized as the root (*Radikalvermögen*) of all our cognition (A 114).

Transcendental apperception is the ultimate subjective condition of cognition for Kant since "[. . .] it is only because I count all perceptions to one consciousness (the original apperception), that I can say of all of them that I am conscious of them" (A 122). And only if I am thus conscious can representations be meaningful

---

[9] Already in the *Dissertation* of 1770, Kant rejects the idea that humans can have intuitions of noumena (Ak 2:413). It is therefore very unlikely that the self of which one is intuitively aware according to R 4675 is the soul, which is an intelligible object.

[10] That they have such a real use seems to be presupposed by all traditional metaphysics. Hence, Kant's problem is how such a presupposition can be legitimized.

[11] Although their interpretations of the argument are radically different, both Paul Guyer in *Kant and the Claims of Knowledge* (New York: Cambridge University Press, 1988) and Wolfgang Carl in *Der Schweigende Kant*: *Die Entwürfe zu einer Deduktion der Kategorien* (Göttingen: Vandenhoeck & Ruprecht, 1989) agree on this point.

[12] On Carl's reading, the unity on which all apperception depends is the ontological unity of the soul qua substance. On Guyer's reading, the unity is the temporal unity, i.e., continuity, of the self. I take issue with both of these interpretations in C. Serck-Hanssen, "Apperception and Deduction in the *Duisburg'sche Nachlass*," *Kant und die Berliner Aufklärung*, vol. 2 (Berlin: de Gruyter, 2001), 59–69.

to me as a cognitive being.[13] However, the unity of mind that is requisite for the consciousness of representations, is not simply given; it requires some special kind of consciousness, for "this unity of consciousness would be impossible if the mind in the cognition of the manifold could not become conscious of the identity of function" (A 108). Moreover, the consciousness of the identity of mind also requires a special kind of consciousness. "For the mind could not think its identity and indeed think it a priori, if it did not have before its eyes the identity of its acts" (*ibid.*).

Recall that in the *Duisburg'sche Nachlass*, apperception in the sense of consciousness of thinking could also be regarded as consciousness of the rules of such an activity. The argument just presented from the A-edition of the first C*ritique* suggests a similar conception of apperception. First, since consciousness of the identity of acts and functions is a condition of the unity and identity of mind, this kind of consciousness must be involved in transcendental apperception *qua* ability to form the 'I'. Second, both of the objects of this consciousness, viz., functions and acts, can be characterized in terms of rules.[14] Before we say more about the relation between rule-consciousness and self-consciousness, we must, however, look at the B-edition's treatment of our topic.

Kant begins § 16 with the well-known claim that it must be *possible* to accompany all my representations with 'I think'. If not, something could be represented in me that could not be thought, which would mean that the representation would be impossible or at least be nothing, i.e., represent nothing to me (B 131–32). Just after the opening lines, he claims that the representation 'I think' is an act of spontaneity, and he announces that he is going to call *it* pure apperception or original apperception, because it is the (kind of) self-consciousness that, *in bringing forth the representation 'I think',* must accompany all other representations (B 132, my emphasis). It is, however, unclear what this *it* is which Kant calls pure apperception. At first it seems to be the representation 'I think'. But in the next passage he claims that apperception is what brings forth the representation 'I think'. Although a bit forced, it is possible, grammatically speaking, to interpret the "it" in the first clause as referring not to the representation 'I think' but to spontaneity. (*Ich nenne sie [die Spontanität] die reine Apperzeption*). But it is also possible that Kant intends "pure apperception" to cover both the representation 'I' or, to be more precise, "I think",

---

[13] Kant does not deny that there are representations that we are not conscious of. Elsewhere he maintains that representations that cannot be accompanied by the "I think" may still exist in us (and in animals) and affect our feelings and behaviour, but as such they would not be anything for me as a *knowing* being. Cf. *Anthropology* at Ak 7:135–137, 397 and letter to M. Herz of May 26th 1789, Ak 11:52. For more on this point see Henry E. Allison, *Kant's Transcendental Idealism*: *An Interpretation and Defence*, revised & enlarged edition (New Haven: Yale University Press, 2004), 163–64.

[14] For the identification of function and act see, e.g., the Clue (A 78–9/B 104–05) and § 20 of the B-deduction (B 143). That acts, or equivalently functions, are to be conceived as rules is clear from Kant's remark that the most appropriate way to characterize the understanding is as the ability of rules (A 126). This characterization is to replace "spontaneity of cognition" and "the ability to think or judge". Earlier in the *Critique* the latter ability was conceived in terms of the different functions of judgment (A 70/B 95).

and the ability to form this representation. As we shall soon see, these two readings are in fact compatible when we add some further qualifications. Kant does indeed hold that spontaneity is an ability that is intrinsically apperceptive, and that it is by virtue of this property that we as thinkers can bring forth the representation 'I think'. As such, spontaneity is later called the "*synthetic* unity of apperception" (B 133, 133n). The 'I' or the 'I think' that is dependent on this synthetic unity, is later referred to as the "analytic unity of apperception" (B 133).

Kant argues that the principle of apperception reveals (*erklärt*) that a synthesis and, indeed, also a consciousness of this synthesis is necessary, since without it that (*jene*) thoroughgoing identity of self-consciousness cannot be thought (B 135).[15] The consciousness of this synthesis must be a consciousness of the synthesis involved in the synthetic unity of apperception which Kant claims is "[...] that highest point, to which we must ascribe all employment of the understanding, even the whole of logic, and conformably therewith, transcendental philosophy. Indeed this ability is the understanding itself" (B 133n). This implies that in the B-edition consciousness of synthesis is part of the concept of original or pure apperception *qua* ability.

According to Kant, a synthesis is necessary, since a consciousness that adheres to each representation but does not adhere to them as unified would have no identity (B 133). In this case, each instance of consciousness would be a distinct consciousness. A helpful analogy would be different thinkers entertaining one thought each (B 407–08). But this cannot be the whole story, so to speak. Different representations and the consciousness that adheres to them could be brought together in a number of different ways without thereby constituting a representation of a single, numerically identical consciousness. Again, the analogy proves helpful. Neither bringing the thinkers together in one room, nor tying them up together would turn their thoughts into one (complex) thought or turn them into one thinker. Hence, in order to do justice to Kant's conception of apperception as the ability to bring forth the 'I', the required synthesis must be something more than a mere aggregation of elements. Moreover, although some synthesis is necessary to avoid a dispersed consciousness, the 'I' Kant seeks to establish contains more than this. Already in § 16, he conceives of it as 'I *think*'. In the later part of the deduction (§ 25) he expands the 'I' even further and maintains that in the synthetic original unity of apperception I am conscious of my capacity to combine (B 158) and of my spontaneity, a consciousness on the basis of which I call myself intelligence (B 157n).

---

[15] A few lines later, Kant seems to withdraw from the claim that awareness of the synthesis is necessary. Patricia Kitcher, therefore, argues that we need not take this claim seriously and that it is a lapse on Kant's part, a so-called epistemologist's fallacy. See Kitcher, *Kant's Transcendental Psychology* (Oxford: Oxford University Press, 1990), 126–27. Guyer also criticises Kant's argument. On his reading, Kant proceeds from the claim that a synthesis is necessary to the claim that there is a necessary synthesis and thus falls victim to the fallacy of necessity. See Paul Guyer, "Kant on Apperception and A Priori Synthesis," *American Philosophical Quarterly* 17(1980), 205–12. But this reading is misguided. The synthesis is a priori, and hence necessary, because it is governed by a priori rules. That is, the synthesis is necessary in the sense of being a necessary condition of cognition.

To get any further in explaining why a synthesis is necessary, we need to distinguish between two different senses of "unity of consciousness", both of which are present in Kant's argument as presented above. On the one hand, unity can be a property of consciousness or mind, viz., that of being one and not many. On the other hand, it can be a property of representations belonging to consciousness, viz., that of being a unity or whole. The two meanings are clearly intimately connected.[16] If three different thinkers each think of one pebble, the thoughts of these three thinkers do not constitute one thought of three pebbles. Only a single thinker, it seems, can have such a unified thought. But rather than being the solution to Kant's problem, the different senses of "unity of consciousness" merely restate it. The reference to a single mind, consciousness, or thinker does not help us, because unless we are told something more about consciousness it is hard to see how the mere presence of three individual thoughts (e.g., 'Jill goes to school', 'Jack goes to school', 'Jane goes to school') in one consciousness could make them a unified whole (viz., 'Jill, Jack and Jane go to school'). The ability to have such single complex representations can, however, be seen as the essential feature of a mind or consciousness. This suggests that the problem of the unity, or singleness, of mind must be approached through a more careful analysis of what is involved in a single complex representation.

Kant characterizes a single complex representation as having its parts combined by "a function of unity". Such a function combines many representations into one (A 69/B 94, A 79/B 104). For instance, mathematical functions act in this way: the sum function combines 7 and 5 into 12. According to Kant, however, the most fundamental functions are the rules of thought or judgment that combine concepts into judgments. With this background, the necessary synthesis in § 16 can be understood not as a mere aggregation of parts but as a combination according to the rules of thought or judgment (Cf. § 19). This makes the point concerning the necessity of a consciousness of synthesis essentially the same as that in the A-deduction (A 108). Both can be formulated as the claim that the mind could only be one and the same in all its thoughts and could only be *conscious* of itself as such, i.e., as an 'I', if it were conscious of the rules of its combination. We are now finally in the position to see why this is the case and why transcendental apperception can be regarded as rule-consciousness.

On the basis of the analysis of a single complex representation, the problem of the unity and identity of consciousness can be raised afresh. Instead of conceiving of consciousness as a kind of object in which a number of representations are somehow contained, it can be conceived as something with the ability to have single complex representations. In accordance with the analysis of what is required for such representations, this ability must be that of running through and combining the elements (representations) in such a way that a unity and not a mere aggregate is obtained (A 77–8/B 102–04, B 129–30). This implies that in moving from one element, A, to the next, B, that which acts while representing B must not only retain A, but also be committed to the way it represented A and must hence be aware of the identity

---

[16] For the reciprocity between the two see also Allison, *Kant's Transcendental Idealism*, 166.

of the function, i.e., the rule for its representing A.[17] In so far as something has the ability to combine representations according to these constraints, it is a single mind. Its unifying ability is, in other words, constitutive of its identity. Whether that which acts (combines) is one or many according to other principles of individuation does not matter. Kant's understanding of the mind and its identity is thus quite opposite to the rationalist tradition. According to the rationalists' view, the objects of mind have a special kind of unity, because they belong to a single and substantial substance, the soul. In Kant's view, however, an object of mind, i.e. a representation, belongs to a single mind if it (the representation) is unified in the proper way.

This account of what constitutes a single mind also contains the solution to the final problem: How is the consciousness of the identity of the mind, expressed by the "I" or "I think", possible? Since a single mind is to be characterized in terms of the ability to form single complex representations, a consciousness of the mind's identity is nothing more than the consciousness of this ability. This explains why Kant claims that through the 'I think', I only represent to myself the spontaneity of my thought, and I exist as an intelligence that is only conscious of its ability to combine (B 157n, B 158). Although the actualization of 'I think' might involve a certain shift of focus, no new kind of consciousness beyond that already involved in the ability to form single complex representations seems to be required in order to account for the possibility of being conscious of the mind's identity in this sense. This is so, because the ability to form single complex representations is the ability to identify, keep track of, and be responsive to the rules through which the different elements are represented.[18] However, to have such an ability is arguably to be at least tacitly aware of having it, since without such awareness, constitutive features of normatively guided action, such as the possible recognition of mistakes and acknowledgement of the commitment to the rules one follows, might just as well be missing. This suggests that the appropriate understanding of this ability (to combine) is that the act of combining according to rules is in a sense inherently reflective.[19] It is thus something more or, rather, other than a blind or mechanical operation; it is not like the planets' movements according to the laws of gravity.

---

[17] This line of thought is only made explicit in the A-deduction (A 103), but it also fits well with the argument in the B-deduction.

[18] One must recall that Kant's analysis is not intended as a factual description of what goes on in the mind, but as a *quid juris*, i.e., a normative analysis of what conditions must be met if cognition, and particularly a priori cognition, is to be possible. For more on this see below.

[19] Karl Ameriks and Dieter Sturma have advocated a similar view. They both raise it as a criticism of Dieter Henrich's claim that it is first with Fichte that this conception of consciousness is found. See Sturma, *Kant über Selbstbewusstsein* (Hildesheim/Zürich/New York: Georg-Olms-Verlag, 1985) and Ameriks, "Kant, Fichte and Apperception," in Karl Ameriks, *Kant and the Fate of Autonomy: Problems in the Appropriation of the Critical Philosophy* (Cambridge: Cambridge University Press), 234–66. See also Robert Pippin, "Kant and the Spontaneity of Mind," *Canadian Journal of Philosophy* 17(1987), 449–75 and Allison, *Kant's Transcendental Idealism*, 170.

So far I have argued that "transcendental", "pure" or "original apperception" is to be interpreted as rule-consciousness.[20] But how are we to understand the kind of consciousness involved? In the *Reflexionen,* Kant identifies apperception with perceptual consciousness, and this idea is repeated at least once in the first *Critique* (A 342/B 401). Given that, one might assume that 'apperception' refers to a mental state in which we perceive special kinds of non-spatial inner objects, namely rules. But this is clearly not Kant's view. First, apperception is a *transcendental condition* of experience in general (*überhaupt*), and as such it is also a condition of any kind of perception, inner or outer. Hence, if apperception were a perceptual state it would become its own condition. Second, Kant emphasizes that to think of transcendental apperception as a kind of inner perception would be to conflate it with inner sense (A 107), and he accuses the different systems of psychology of having committed just this kind of mistake (B 152).[21] Third, if apperception were a mental or psychological state, 'apperception' would be a psychological concept whose content would have to be established empirically. But the content of 'apperception' is established in quite a different way, viz., by showing how such consciousness can give rise to the thoroughgoing identity of apperception, which Kant takes to be a normative condition for thought and cognition. 'Apperception' is a concept of a standard that anything that has the right to be called thought must live up to, it does not concern how an empirical mental state or process can be described. As such, apperception need not be something experienced at all. And Kant goes further than this. In R 5661, he explicitly denies that consciousness of thinking is an experience (Ak 18:319). But if apperception is not a psychological state, what is it?

As we have seen, Kant's approach to apperception centres on the analysis of functions of thought and cognition. This suggests, of course, that rather than being a concept of a particular kind of mental state, 'apperception' is a functional concept. Apparently this also fits well with Kant's argument in the Paralogisms: the principle of apperception does not yield any knowledge of that which thinks.[22] On the functionalist's reading, Kant's point is simply that no metaphysical conclusions can be drawn from a functional description of thought. But although Kant would surely agree that such a step would be unwarranted, his point is a different one. The error of speculative (rational) psychology according to him is not primarily that it aims at establishing the kinds of entities that "do the thinking" instead of focusing on the cognitive tasks that must be performed. The fundamental error is, rather, the hidden conceptual shift from the *determining* to the *determinable* I (B 157n). In our attempt to cognize the former, we unavoidably change our perspective and illicitly present what is essentially subjective and active as an object and, thus, conflate thinking with things (B 411n). As he says in the A-edition "This 'I' is as little an intuition

---

[20] In the present context I take these three to be equivalent. Unless otherwise indicated I shall henceforth use "apperception" as shorthand for all of them.

[21] Cf. also *Anthropology,* Ak 7:142. For Kant's rejection of the view that his transcendental philosophy is a kind of empirical psychology cf. also R 4900 and R 4901.

[22] For such a reading of Kant see, e.g., Patricia Kitcher, *Kant's Transcendental Psychology*, and Andrew Brook, *Kant and the Mind* (Cambridge: Cambridge University Press, 1994).

as it is a concept of any object, but the mere form of consciousness" (A 381). The same point is emphasized in the *Anthropologie in pragmatischer Hinsicht*: "The I is neither an intuition nor a concept, it is not a determination of an object at all, but an act of understanding of the determining subject in general, its consciousness of itself, pure apperception, therefore belongs merely to logic" (Ak 7:397).[23]

The idea that functions are inextricably related to subjectivity and activity has no place in contemporary functionalism.[24] The main reason for dismissing a functionalistic reading of Kant is, however, that, while functionalism attempts to reduce "consciousness" to concepts of function, Kant holds that the functions or, equivalently, rules, are things of which the epistemic subject must be conscious. But this just brings us back to the problem with which we started: "What kind of consciousness is involved in apperception?"

We have already seen that 'apperception' should not be taken to be a concept of a kind of psychological state. Given the peculiar status Kant assigns to the functions or rules of thought, that they are to be the conditions of all objecthood, we realize that not even they can be conceived using an object as a model. In this sense, the essential subjectivity of apperception also applies to the rules involved in apperception. Following this line of thought, apperception does not consist of some relation between two independent entities: consciousness and a rule. 'Apperception' is the concept of a kind of consciousness that is always already intertwined with the rule. The recognition of the rule as the same rule must somehow take place directly and not require the acts and concepts on which object recognition depends.[25] More fundamentally, the identity of the rule *qua* rule, its normative force, is not something it simply has but is constituted by it being taken or grasped as such. Apperception for Kant thus consists in this immediate and internal relation between consciousness and normativity.

As such a kind of consciousness, apperception can be seen to be present not only in thought and cognition but in all human activity (spontaneity), and it manifests itself in certain forms of responsibility. In the theoretical and practical[26] domains, this responsibility expresses itself in our practice of reason giving. Reason giving is

---

[23] See also *Anthropology,* Ak 7:141, where Kant says that the character of the intellectual power of cognition is the spontaneity of apperception, i.e., the pure consciousness of the act that constitutes thought, and as such it belongs to logic.

[24] Indeed it is those who criticize reductive theories of consciousness who often appeal to the irreducibility of subjectivity. Cf., e.g., Nagel "What Is It Like To Be a Bat?" and F. Jackson "What Mary Didn't Know," in D. Rosenthal (ed.), *The Nature of Mind* (New York: Oxford University Press, 1991).

[25] This implies that apperception involves the capacity of reflective judgment, which is heautonomous, i.e., it gives the rule to itself (*KU*, Ak 20:225). Reflective judgment is also argued to schematize without a concept, i.e., it is a capacity that finds rules without already presupposing them (*KU*, Ak 4:287). The doctrine of reflective judgment is only worked out in the last *Critique*, but as has been emphasized by Longuenesse, it is already assumed in the first. Cf. B. Longuenesse, *Kant and the Capacity to Judge: Sensibility and Discursivity in the Transcendental Analytic of the Critique of Pure Reason* (Princeton: Princeton University Press, 1998), 163–66.

[26] Cf. A 546/B 574 where Kant relates apperception to understanding, reason, and, finally, to practical reason.

not a practice independent of thinking, believing, or acting; rather, these activities only take place under the consciousness of an ought (*Sollen*). Also, in the aesthetic domain the consciousness of a norm is presupposed. I take pleasure in the object, because in apprehending it I find myself living up to the standard of how I and everybody else ought to respond to it (KU § 22, Ak 5:239).

8.2 Inner Sense

In the first *Critique* and later works, Kant carefully distinguishes between appercep-tion and inner sense. The distinction is perhaps put most clearly in the *Anthropology* where inner sense is said to be the consciousness of what a human being undergoes, as opposed to apperception, which is the consciousness of what such a being does (Ak 7:161). Note also that inner sense is precisely a sense, i.e., it is a sensible faculty whose representations are classified as intuitions. In earlier works, the picture is more muddled, but that need not concern us here.[27]

According to Kant's critical epistemology we have no knowledge of objects as they are in themselves beyond space and time. Since, traditionally, the self has been seen as a special kind of object, viz., the soul, to which one has privileged and infallible epistemic access, it is of utmost importance to him to show that this is not the case. In the Dialectic of the first *Critique* he writes: "[. . .] a great and indeed the only stumbling-block or unanswerable objection to our whole critique would be if there were a possibility of proving a priori that all thinking beings are in themselves simple substances" (B 409). His doctrine of inner sense must be understood against this background. It aims at showing that we have no epistemic access to ourselves as we are in ourselves and that we only know ourselves as appearances. This constraint is in itself, however, insufficient for understanding Kant's theory, which is partly presented in confusing and extremely cryptic terms.

Kant's theory of inner sense involves four central claims. The first concerns the object of inner sense. Inner sense is the ability to be intuitively aware of the self, soul, or mind and its inner states (A 33–4/B 49–50, A 342/B 400). The second claim is that time is the form of inner sense (A 33/B 49). All inner objects are immediately in time, while all outer objects are immediately in space and only mediately in time. The third claim is that we only know ourselves through inner sense insofar as we affect ourselves (B 67, 69, 153). The fourth claim is that inner sense has no manifold of its own; the representations of outer sense are what really make up its material (B 67).

---

[27] According to *Novo Dilucidatio*, inner sense is the faculty through which we are aware of being inwardly affected by real existing bodies (Ak 1:411). The same conception of inner sense seems to be presupposed in *Von dem ersten Grunde des Unterschieds der Gegenden im Raume*, since the intuitions of inner sense are claimed to prove the reality of space (Ak 2:373). In *De mundi sensibilis atque intelligibilis forma et principiis* (1770) inner sense is said to contain the phenomena to be investigated by empirical psychology (Ak 2:397) and is the ability to intuit representational states (Ak 2:405). In *Die falsche Spitzfindigkeit* 'inner sense' shows close affinity to what he later refers to as apperception (cf. Ak 2:59).

The first puzzle is that the claim, that the object of inner sense is the soul and its states, appears to conflict both with the claim that all the material of inner sense belongs to outer sense and with the often repeated claim that we have no intuition of the soul. To solve this problem, we first need to draw a distinction between "representation" in the sense of a mental state and in the sense of the content of such a state. If the objects of inner sense are the mental states themselves (or even the soul), how can the material of inner sense still come from outer sense? Certain points in the *Anthropology*[28] suggest a way out of this puzzle. What Kant seems to have in mind is that we can only *represent* mental states to ourselves by means of analogies taken from outer sense.[29] Notice also that in the same work Kant qualifies his claim about the soul being the object of inner sense. Here he maintains that in inner experience the human being *believes* (*glaubt*) that she perceives a soul in herself (Ak 7:161). This suggests that although we have no genuine intuition of a soul, the mental states we represent to ourselves through spatial analogies are neither represented as free floating, nor simply as attached to the 'I' of apperception. They are, rather, represented *as if* they were states of a soul. Somewhat cautiously, representations of these states could therefore be called representations *of* the soul and its states.[30]

That both inner and outer objects are represented with spatial properties certainly does not imply that Kant identifies them. To see what distinguishes inner objects and inner knowledge from outer objects and knowledge thereof, we now need to turn to the other aspects of Kant's theory of inner sense.

As we have seen, the ascription of representations to oneself as the identical subject of thought is not a function of inner sense but of apperception. Hence, what the theory of inner sense is supposed to explain cannot be this. Rather, the aim of this theory is to account for another distinction, one that needs to be made within the class of all representations that I am able to relate to myself through apperception. While apperception is a condition of all possible experience, the special task of the theory of inner sense is to explain how an experience of my own experiencing or representing is possible. This does not follow directly from apperception. It is one thing to be able to reflect that the representation of a tree is my representation, i.e., to be able to attach 'I think' to it; it is quite another thing to be able to experience the representing of the tree as something that takes place "in me".

This is where Kant's doctrine of time as the form of inner sense enters into the picture. Already in the *Inaugural Dissertation,* Kant ties the experience of objects

---

[28] Kant claims, e.g., that the causal properties of blood function as an analogy for the temperament of the soul (Ak 7:286–87).

[29] Apart from the general idea that objects of outer sense must be employed as analogies, Kant never tells us how we can make use of such analogies to identify and distinguish between our different mental states. If the theory is taken as a contribution to empirical psychology it is surely a deficiency. Kant's theory of inner sense is, however, aimed at a more fundamental level of inquiry, namely inner experience in general (*überhaupt*). As such, it is supposed to supply only the general conditions of experiencing mental states and their content.

[30] Note that such representations cannot support any knowledge claims about the soul as an entity of transcendent metaphysics.

(*Gegenstände*) to space and the experience of states, representational states in particular, to time (Ak 2:405). A similar point is made in the discussion of time in the *Critique of Pure Reason* (A 34/B 50). Kant can hardly be said to have given any explanation of why this is the case, but the following line of thought seems to fit well with his views.[31]

In so far as objects are considered only from the point of view of their spatial qualities, they do not justify the application of temporal predicates. It is senseless to ask when the car is, to say that the house is five minutes long, or that the rock is earlier than the tree. In order for temporal predicates to be applicable to spatial objects, a certain ontological shift is needed. It is only if we consider spatial objects from the point of view of their existence, e.g., as events or processes, that they fall within the scope of temporal predicates. I can determine the age of the house when it is seen as something that comes into existence through a building process, and I can make claims about the tree being younger than the rock if I take both of them to be events in the history of nature. In this reading, outer objects are only mediately temporal for Kant, since it is only their spatial properties that are immediately given to us in intuition. Their temporal properties only apply to them through an ontological shift.

When we turn to inner objects, the picture is different. Although inner representations are represented as spatial, they do not, of course, really have spatial properties. This holds for both senses of the term representation. My mental state of seeing a Chinese wall has no length; neither does a line I imagine when I conceive of it as a mental object. But according to Kant all inner objects have temporal properties. This is supposedly true not only of mental states such as thoughts, perceptions, and memories, which clearly begin and cease to exist at determinate points in time, but also of the way we experience the content of such states through inner sense. If I am to experience myself as representing (seeing or imagining) a line, I cannot experience the line as I do in outer intuition, viz., as a given spatial whole. According to Kant, I must experience the line as being generated through a chain of events in which the different segments of the line are produced in me successively (B 154). Moreover, inner objects (both states and their content) do not only appear in time. They are necessarily temporal, because time is the a priori form of their appearing. Inner objects are in this sense *immediately* temporal; if they appear at all, they appear with temporal properties. In other words, as opposed to outer objects no ontological shift is required to make inner objects temporal.

The immediate temporal character of inner objects does not, however, imply that these properties are derived from inner sense alone. Inner sense is that *to which* inner objects are given, but not that *by means of which* they are given. For the conditions of the givenness of inner objects, we need to turn to Kant's theory of self-affection.

The doctrine of self-affection can be seen to contain a general and a more specific point. The general point follows from the transcendental deduction in which Kant argues that the manifold of our sensible forms must be subject to what he calls the

---

[31] In the following section, I draw on Allison, *Kant's Transcendental Idealism*, (N.B. first edition), 255ff.

transcendental synthesis of the power of imagination (B 151–52). Since these forms pertain to us, this subjection must be seen as a kind of self-affection. Moreover, since *every* object that can be known by us presupposes this kind of synthesis it, of course, follows that "I know myself, like other phenomena, only as I appear to myself, not as I am to the understanding" (B 155).

When it comes to Kant's views on inner sense it is, however, the specific point about self-affection that is of interest. This thesis goes further than the general point. It concerns not only how a given manifold must be apprehended or interpreted if it is to be known but also the condition of the (inner) manifold being given, i.e., the condition of its possibility as *intuition*. Consider the following claims: "When the ability to be aware of itself is to seek out (apprehend) that which lies in the mind, it must affect the mind and only in this way can it bring forth *an intuition* of itself" (B 68, amended translation, my emphasis).[32] Moreover, "we only *intuit* ourselves as we are inwardly affected" (B 153, my emphasis). What then does this kind of self-affection amount to?

As opposed to inner sense, self-affection is not a receptive ability but an activity that requires conceptualization. Kant tries to illustrate his point the following way:

> This we can always perceive in ourselves. We cannot think a line without drawing it, or a circle without describing it, we cannot represent the three dimensions of space save by setting three lines at right angles to one another from the same point [. . .] The understanding does not therefore find in inner sense such a combination of the manifold but produces it, in that it affects that sense (B 154).

Despite Kant's way of expressing it, his aim is not to establish what is required in order to think or imagine an object. As the first clause and context (the problem of inner experience) indicate, the argument concerns the second level problem of how we can represent ourselves (to ourselves) as things in which representations occur. Kant attempts to draw our attention to what he takes to be the significant mark of such inner experience. I do not only represent my representational state, e.g., thinking, as an event or process in time. Also the inner *object*, e.g. a line, is necessarily represented as successively appearing through a process in time, as opposed to how I would represent the object of such a state if I were to refer it to an object outside me, e.g., to the cover of the book in front of me. The crucial point is, in other words, that all objects of inner experience are represented as events. According to Kant's argument in the *Analogies,* we can have no experience of events or processes without the concept of causality (A 211/B 255–56). Only through this concept do we achieve an awareness of the irreversible time order that characterizes the experience of events: what happens necessarily happens after a time during which it was not yet present.[33] Hence, the specific point about self-affection is that there can be no intuitions of inner sense without the concept of causality.

---

[32] The important difference between my reading and Allison's is that he takes self-affection to be involved only in the *determination* of inner intuition, while I argue that it is a condition of the very intuition itself.

[33] For an insightful reading of Kant's views on causality see Allison, *Kant's Transcendental Idealism*, 246–60.

That the concept of causality is constitutive for inner intuition shows not only that there is an important difference between inner and outer sense and the corresponding experience; it also marks the crucial difference between inner sense and apperception. In inner sense I am aware of being a soul or mind in which a process of representing takes place. In apperception I am aware of myself as a spontaneous subject that represents according to rules.

8.3 Kant's Conception of the Mind–Body Problem

Just as there are two concepts of consciousness, 'inner sense' and 'apperception', there are two sets of relations between consciousness and nature for Kant. On the one hand, there is the relation between the mental and the physical, i.e., inner and outer nature; on the other, there is the relation between our apperceptive spontaneity and the whole of nature (inner and outer). The two different relations give rise to two different problems. The first problem is how the mental and the physical can affect each other. The second is how spontaneity, which falls outside the domain of causal laws and nature altogether, can bring about an event in nature.[34] Here I can only give a brief sketch of Kant's way of dealing with the two.

8.3.1 The Mental and the Physical

From Kant's theory of inner sense, it follows that we know a priori that every mental object has some cause. Without this assumption, no representations of inner states and objects would be possible. Moreover, the claim that inner sense has no manifold of its own appears to be a claim not only about the form, i.e., its spatiality, but also about the origin of this content. One of the argumentative lines in the Refutation of Idealism (B 276–77n), namely that space could not be a product of imagination, indeed turns on this reading of the claim.[35] If so, Kant also holds that at least some mental events must be conceived as perceptual events caused by physical events. Finally, in his moral reasoning Kant argues for a strict determinism with respect to agency: "[i]f we could exhaustively investigate all the appearances of men's will, there would not be found a single human action which we could not predict with certainty and recognize as proceeding necessarily from its antecedent conditions" (A 549–50/B 577–78).[36] Kant also holds that where there is a cause there is a law (B 232). Hence, his metaphysical position seems simply to be that there is genuine causal interaction between the mental and the physical and that in addition to physical laws there are psychological and psycho-physical laws.

---

[34] Since spontaneity is free according to Kant, the question is never how nature can affect spontaneity.

[35] Cf. Allison, *Transcendental Idealism*, 293–95.

[36] See Henry E. Allison, *Kant's Theory of Freedom* (Cambridge: Cambridge University Press 1990), 32–4.

This view seems, however, to be challenged by certain claims in the *Metaphysical Foundations of Natural Science.* Here, Kant argues that psychology can never become a science proper because we can find no strict laws that govern inner objects (Ak 4:471). How can these apparently conflicting views on the nomological character of the mental be reconciled?

To reconcile Kant's thoroughgoing empirical determinism with his pessimism concerning the prospects of psychology, we must recall that his account of mental objects takes place at a transcendental level of inquiry. At this level, nothing can be established with respect to the possibility of finding specific empirical laws; however, this is what empirical science requires. In other words, despite what has been established at the transcendental level, the objects of nature (inner as well as outer) may still appear to us without the kind of law-like regularities that are needed to detect, e.g., causes and kinds.[37] Hence, the theory of inner sense only establishes a priori that every inner object must have *some* cause, and that some of these causes must be physical. It follows for Kant that some mental and some psycho-physical laws must exist. The theory of inner sense, therefore, warrants our search for such causes and laws. Nevertheless, it does not guarantee that we will ever succeed in finding them.

Kant's view of the status of psychology is, however, not only fully compatible with his theory of inner sense: two of his arguments can be seen to follow from this theory, or so I shall argue.[38] According to Kant, the first obstacle for psychology becoming a science proper is that only the law of constancy could apply to its objects (Ak 4:471). But how are we to understand this problem? If there is a mathematical law that applies to the entities of psychology, why don't we have what is required to give psychology a scientific status? And what sense are we to make of his analogy: Compared to what is achieved by applying mathematics in the study of bodies, the law of constancy would be an expansion of knowledge that is somewhat (*ohngefär*) comparable to what the study of the properties of the straight line contributes to the whole of geometry (loc.cit.). The straightforward reading of the analogy is that Kant takes the object of inner sense to be richer than what can be captured by applying mathematics (the law of constancy); such an approach would only inform us about a certain part or aspect of our object. But again, why would this partial knowledge not qualify as scientific? Why should a scientifically minded psychologist indeed not to strive to limit her object accordingly? I my view, however, Kant's theory of inner sense suggests quite a different reading of the analogy.

Psychology is the science of the objects of inner sense. Nevertheless, the objects of inner sense do not constitute an independent domain of study, since according to Kant inner sense has no material of its own. As I have already suggested, this is not only a theory about the form of this material, viz., its spatiality and its dependence on conceptualization, but also about its origin. If so, a genuine psychology would, at the

---

[37] Cf. *Ibid.*, 34.

[38] Notice that the details of this theory only appear in the B-edition of 1787, two years after the *Metaphysical Foundations.* If my reading is correct, he must, however, already have worked out the main ideas of this theory before 1785.

very least, have to take some representational states to be events caused by external occurrences, and, most likely, perceptual events ought to play a fundamental explanatory role. A psychology that purports to restrict itself to the study of a "purely inner" content would be somewhat comparable to a geometry that restricts itself to the study of properties of the straight line, because, just like such a geometry, it would have misconceived the scope of its own science. Hence, the real problem is not that the application of the law of constancy tells us very little about the purely inner but that it fails to say anything about the non-eliminable relation between inner and outer objects. As such, the first obstacle is only an obstacle to a certain, and in Kant's views, misconceived understanding of psychology.[39] As we shall now see, the second obstacle goes further than this.

The second obstacle for psychology is that inner objects can only be individuated and separated in thought; they cannot be stored independently of thought and cannot be randomly conjoined (Ak 4:471). Although Kant does not work out his argument, his position seems to be that with respect to inner objects there is, in some important sense, no concept-independent truth of the matter. Why he would hold such a view is not clear unless we turn to his theory of inner sense. From this theory, we know that inner objects are not simply given to inner sense but that they presuppose an activity of conceptualizing them as events. Outer, i.e., physical objects are, however, represented as extended and separate prior to any such conceptualization of them.[40] Since the concept of causality is constitutive of inner objects and only applies to outer objects in so far as they are already individuated in space, causal claims in the two fields are treated differently. In the physical field, they can be justified by experiments with objects that can be identified independently of any causal claims; in the psychological field, they cannot. For Kant, this implies that neither strict mental laws, nor strict psycho-physical laws can ever be established on the empirical level and hence that psychology cannot provide more than natural descriptions of the soul (Ak 4:471). But Kant still believes that from the transcendental point of view both the mental and the physical are determined by laws and hence that no freedom is possible in nature at all.[41]

---

[39] In "Kant on Psychology: How Not to Investigate the Human Mind," in Eric Watkins (ed.), *Kant and the Sciences* (New York: Oxford University Press, 2001), 178, Thomas Sturm argues that the arguments in *MAN* are only aimed at psychology conceived as a purely introspective science. I agree that this is the case with respect to the first obstacle discussed above, but not with the second and third. The third obstacle is not, however, discussed in this chapter.

[40] Notice that in the Analogies, Kant claims that the causal principle is only regulative and not constitutive, since it is not necessary for the mere possibility of appearances (A 179–80/B 222–23). I take this point to hold only for outer appearance, since their mere possibility is not contingent on them being events. As is clear from the above reading of Kant's theory of inner sense, the same does not hold for inner objects or appearances.

[41] Although Kant's theory of inner sense has some interesting similarities with Davidson's views on the mental, they differ on the important issue of anomaly and freedom. That mental laws cannot be scientifically justified does not imply anomaly in Kant's view. Accordingly, he argues that there is no freedom in anything that belongs to the empirical (A 549–50/B 577–78).

8.3.2 Spontaneity and Nature

The problem of the relation between spontaneity and nature is dealt with in all three *Critiques.* In the first *Critique*, Kant argues for the logical possibility that although all events in the sensible world take place in accordance with laws of nature it could still be the case that they are grounded in some other kind of causality, that of freedom (A 536/B 564). His solution depends on transcendental idealism, the doctrine that space and time only apply to appearances and not to things in themselves. This distinction between appearances and things in themselves also applies to the subject and allows Kant to ascribe two different characters to it. From the point of view of the subject's empirical character, "[. . .] its actions stand in thoroughgoing connection with other appearances in accordance with unvarying laws of nature" (A 539/B 567). But "we should also have to allow the subject an *intelligible* character, by which it is indeed the cause of these same actions as appearances, but which does not itself stand under any conditions of sensibility, and is not itself appearance" (*ibid.*).

The first *Critique* thus attempts to show that it is *conceivable* that *qua* the human being's intelligible character, spontaneity is the ground of her actions as appearances.[42] It also suggests how a stronger claim could be made, namely that we *must* conceive of ourselves in this way:

> Reason does not here [in willing] follow the order of things as they present themselves in appearance, but frames for itself with perfect spontaneity an order of its own according to ideas, to which it adapts the empirical conditions, and according to which it declares actions to be necessary [. . .] And at the same time reason also presupposes that it can have causality in regards to all these actions, since otherwise no empirical effects could be expected from its ideas (A 549/B576).

We cannot but think of ourselves as beings that have the capacity to begin absolutely and of ourselves a series of events in the empirical world, because this is the only way we can make sense of claims concerning what we *ought* to do (A 547/B 575) and the only ground on which we can make sense of our practice of blaming and praising each other (A554–55/B 582–83).

This second step beyond mere conceivability is not worked out until the two next *Critiques.* In the *Critique of Practical Reason*, Kant argues that the moral law and freedom are reciprocal concepts (Ak 5:29). To get from this conceptual relation to a substantive conclusion, he argues that the moral law is a fact of reason (Ak 5:31).[43] Since the moral law appears to us as a categorical imperative that demands unconditionally, we must consider ourselves as free in the positive sense of being able to legislate our will. Pure reason must be practical in and of itself or else the unconditional *ought* of the categorical imperative would be senseless. By moving to considerations in the moral domain, we have thus "widened our knowledge beyond

---

[42] I use "ground" instead of "cause" since strictly speaking "cause" only applies to appearances.

[43] Kant also argues for the reciprocity between freedom and the moral law in the *Groundwork* (Ak 4:450), but there he uses the reverse strategy and argues from our conception of ourselves as free to the reality of the moral law (Ak 4:453).

the limits of the world of sense" (Ak 5:50). Even if the reality of spontaneity's causality cannot be theoretically proven, it can be established from the practical point of view whereby its *practical* reality is proven (Ak 5:56).

Kant's view on the relationship between spontaneity and nature culminates in the *Critique of Judgment*.

> In this ability [taste], judgment does not find itself subjected to a heteronomy of laws [. . .] it legislates to itself, just as reason does regarding the power of desire. And because the subject has this possibility within him, while outside there is also the possibility that nature will harmonize with it, judgment finds itself referred to something that is both in the subject himself and outside him, something that is neither nature nor freedom and yet is linked with the basis of freedom, the supersensible, in which the theoretical and the practical power are in an unknown manner combined and joined into a unity (Ak 5:353).

Although we can never know how it is possible that spontaneity begins a series of events in nature, the experience of beauty bears witness to nature being amenable to spontaneity. In the experience of beauty the human being thus finds herself to be at home in the world.

Chapter 9
Physiognomy as Science and Art

Martina Reuter

Abstract The chapter examines Johann Caspar Lavater's studies of physiognomy. Lavater attempted to systematize principles on which we could detect human character traits and dispositions by observing the physical appearance of persons. His optimism regarding the possibility to "read" a person's character from the person's bodily features was based on his understanding of the individual as a unique whole. Lavater combined this holism, anticipatory of Romanticism, with an Enlightenment ideal of systematization and geometrical exactness. The chapter focuses on three aspects of Lavater's conceptual framework: his view of the soul, the relation between the particular and the universal, and the relation between necessity and accidentiality, including the possibility of freedom. It is claimed that Lavater's attempt to conceptualise the relation between the particular and the universal interestingly resembles the view later systematised by Immanuel Kant in his *Critique of the Power of Judgment* (1790). A comparison with Kantian aesthetics will illuminate more closely why and how Lavater's theory is pseudo-scientific. When considering the relation between necessity and freedom, Lavater defends the view that necessity is inherent in the constitution of a person's physiognomy, and he takes a critical distance from the Enlightenment belief in the power of education. The chapter concludes by examining to what extent Lavater's physiognomical theory influenced the thought of the early feminist writer Mary Wollstonecraft. It is claimed that Wollstonecraft emphasises the role of education also in relation to bodily constitution and questions Lavater's preference of necessity over freedom.

Keywords Physiognomy · human character · mind–body relation · Lavater · Kant · Wollstonecraft

The physiognomy developed by the Swiss Protestant pastor Johann Caspar Lavater (1741–1801) is most often depicted as an example of pseudo-science. This reputation is not without grounds. Lavater's attempts to determine a person's character on the basis of his or her facial features remain without any systematic

M. Reuter
Department of Systematic Theology, Academy of Finland and University of Helsinki

justification. Still, it seems that the mixture of horror and ridicule with which physiognomy is rejected bears witness not only to its pseudo-scientific nature as such but to a contemporary trauma, caused by the genocide of the twentieth century, when pseudo-scientific connections between bodily features and character became part of anti-Semitism and other racist ideologies. When evaluating Lavater's work, it is important to keep two things in mind: Racist developments are not a necessary outcome of the idea that there is a connection between anatomy and character, and, even more important, our predecessors had different and in some respects more variable notions of the mind–body connections than we do.[1]

Lavater's physiognomical treatise was a real bestseller of its time, and it appeared in numerous translations, editions, and abridgements. Lavater himself was an international celebrity, who personally knew many foremost contemporary intellectuals. He corresponded on theological matters with Immanuel Kant in the 1770s and Kant's late work on psychological topics, *Anthropology from a Pragmatic Point of View* (1798), is influenced by physiognomical ideas about the relation between bodily features and character.[2] J.W. Goethe knew Lavater well. In his autobiography Goethe comments at length, enthusiastically as well as critically, on Lavater's theology and studies of human character.[3]

In the 1790s, Lavater's work became popular in England, and by the end of the decade seven English translations in several versions had been published, many of them original works in their own right.[4] Among Lavater's enthusiastic readers was the early feminist writer Mary Wollstonecraft (1759–1797)[5] who began a translation of his treatise in the late 1780s.[6] During this period, she also included

---

[1] Lavater's physiognomy is a forerunner to, but must not be confused with, Franz Josef Gall's phrenology, which is a truly scientist attempt to determine character by the form of the human skull and especially the forehead; see Christopher Rivers, *Face Value: Physiognomical Thought and the Legible Body in Marivaux, Lavater, Balzac, Gautier, and Zola* (Madison: The University of Wisconsin Press, 1994), 231 n. 30; and G.P. Brooks and R.W. Johnson, "Johann Caspar Lavater's *Essays on Physiognomy*," *Psychological Reports* 46 (1980), 16ff.

[2] In a letter to Lavater from 1775, Kant emphasises his critical stance on all kinds of religious seremonies, see Allen W. Wood, "Rational Theology, Moral Faith, and Religion," in Paul Guyer (ed.), *The Cambridge Companion to Kant* (Cambridge: Cambridge University Press, 1992), 396.

[3] Johann Wolfgang Goethe, *Dichtung und Wahrheit*, ed. Klaus-Detlef Müller, Goethe Werke vol. 5 (Darmstadt: Wissenschaftliche Buchgesellschaft, 1998), 544–58, 674–85.

[4] See John Graham, "Lavater's *Physiognomy* in England," *Journal of the History of Ideas* 22:4 (1961), 562.

[5] Wollstonecraft is best known for her political and early feminist pamphlets *A Vindication of the Rights of Men* (1790) and *A Vindication on the Rights of Woman* (1792), but she also wrote a handbook on the education of girls, a history of the French revolution, two novels, travel letters, and she translated, collected, and wrote moral stories for children. In addition to these books, she produced approximately 420 book reviews and hundreds of private letters. None of these texts are philosophical or scientific treatises, but they do attribute to Wollstonecraft the title of Enlightenment *philosophe* and an early representative of Romanticism.

[6] Wollstonecraft soon abandoned the translation, which probably was taken up and completed by Thomas Holcroft. According to Wollstonecraft's husband and biographer William Godwin, she translated Lavater's book from a French translation, approved by the author, not from the German original *Physiognomishe Fragmente, zur Beförderung der Menschenerkenntnis und Menschenliebe* (1775); see William Godwin, *Memoirs of the Author of "The Rights of Woman,"* ed. Richard

some of his aphorisms in an edited volume *The Female Reader* (1789) and wrote a posthumously published fragment of fiction "The Cave of Fancy", where she appropriated physiognomical ideas. In 1792, Wollstonecraft published a short but harsh review of an English abridgement of Lavater's treatise. The harsh tone is not directed against Lavater but against the translator-editor Samuel Shaw, whom she accuses of plagiarism as well as distortions.[7] The substantial critique focuses on the subtitle Shaw has given the abridgement: the "Analogy between the Conformation of the Features, and the ruling Passions of the Mind."[8] Wollstonecraft claims that Lavater's physiognomy is not about the analogy between features and passions, a topic Lavater calls pathognomy, but about the relation between the "original formation of the bones" and the powers and inclinations of man.[9] Her focus on "formation of the bones" limits the theoretical range of Lavater's physiognomy but fits well together with her use of his ideas in "The Cave of Fancy". I will show that Wollstonecraft's treatment of Lavater's work is not only critical, but also constructive.

Lavater intends to systematize principles according to which the observer can determine human character traits and dispositions from physical appearance. He attempts to merge the Enlightenment ideal of scientific reasoning about nature with a conception of nature as an animated organic whole. The ideal of systematisation and geometrical exactness is combined with an emphasis on unique particular wholes. In

---

Holmes (London: Penguin Books, 1987), 226. This claim has been repeated by modern commentators; see Rivers, *Face Value*, 68. Still, the claim seems odd: in his preface, Thomas Holcroft refers explicitly to the task of translating from German, and at least once he provides the original German phrase in a footnote. Lavater's quotations from German authors are translated in the text without giving the German originals, while occasional quotes from French authors are provided in French and translated in footnotes; see Johann Caspar Lavater, *Essays on Physiognomy; For the Promotion of the Knowledge and the Love of Mankind*, trans. Thomas Holcroft (London: G. G. J. & J. Robinson, 1789), vi, 113, 157. Based on information from the *English Short Title Catalogue. Eighteenth Century Collections Online*. Gale Group. http://galenet.galegroup.com/servlet/ECCO. In light of these facts, it seems that if it is true that Holcroft completed a translation begun by Wollstonecraft, it is most likely that this was a translation from German. Most translations Wollstonecraft made at that time were from German, not French, which she learned only later, during her stay in France in 1793–1794.

[7] Shaw is accused of plagiarising Holcroft's translation. The review begins by claiming that this is "one of those surreptitious publications, by means of which authors, or their representatives, are deprived of their legal right, to gratify the rapacity of needy book-makers, or the impatience of frugal readers;" see Mary Wollstonecraft, "Contributions to the *Analytical Review* 1788–1797," in *The Works of Mary Wollstonecraft*, vol. 7, eds. Janet Todd and Marilyn Butler (London: William Pickering, 1989), 450. The attribution of reviews published in the *Analytical Review* is complicated, because all reviewers signed only with initials, usually several different, which were not based on their own names. There has been some scholarly debate as to which initials should be ascribed to Wollstonecraft. Janet Todd summarises the debate in her preface to the selection of contributions published in *The Works of Mary Wollstonecraft*. Using stylistic as well as biographical criteria, Todd decides for an inclusive selection and ascribes all contributions signed "M", "W", and "T" to Wollstonecraft. See Janet Todd, "Prefatory Note," in *The Works of Mary Wollstonecraft*, vol. 7, 14–8. It is more or less beyond doubt that the review of Shaw's abridgment is written by Wollstonecraft.

[8] *Ibid.*, 449.

[9] *Ibid.*, 450.

my explication, I will leave the evident problems of empirical justification aside and focus on three aspects of Lavater's conceptual framework: his view of the soul, the relation between the particular and the universal, and the relation between necessity and accidentiality, including the possibility of freedom.

Viewed against the history of conceptions of the soul, Lavater's tripartite model is interesting in its adoption of the Platonist distinction between rational, spirited, and appetitive parts. This fundamentally moral model has historically been less common than the basic Aristotelian psychological model. Simultaneously, Lavater is explicitly Anti-Platonist in his emphasis on the priority of the particular over the universal. In his attempt to conceptualise the relation between the particular and the universal, Lavater makes many remarks which interestingly resemble the view later systematised by Immanuel Kant's aesthetics. A comparison with Kant's view will illuminate more closely in what sense Lavater's theory is pseudo-scientific. Finally, when considering the relation between necessity and freedom, Lavater defends that necessity is inherent in the constitution of a person's physiognomy and takes a critical distance from the Enlightenment belief in the power of education.

I will conclude by examining the similarities and differences between Lavater's physiognomy and Wollstonecraft's appropriation of it. The main difference concerns exactly the role of education. While comparing these two perspectives, I will apply a considerably modified principle of charity. Instead of aiming at an interpretation that is as free of contradictions as possible, I will focus on the tensions and contradictions in their views and aim at making philosophical sense of the origin of these tensions and of what the authors try to express.

9.1 The Physiognomy of Wholes

A spirit of Romanticism is expressed right at the beginning of *Essays on Physiognomy*, where Lavater quotes at length Johann Gottfried Herder's poetical description of man as the image of God. The combination of variety and unity is specific to God's creation as a whole and especially to man created as an image of God:

> God of perfection! How supremely how benevolently hast thou displayed thyself in man! – Behold the human body! That fair investiture of all that is most beauteous – Unity in variety! Variety in unity! How are they there displayed in their very essence! – What elegance, what propriety, what symmetry through all the forms, all the members! How imperceptible, how infinite, are the gradations that constitute this beauteous whole![10]

This quote reveals the promise as well as the most essential restriction of physiognomy, as they are understood by Lavater. The symmetry of forms indicates the promise that the visible symmetry includes traits of character as well as body.

---

[10] Lavater quotes from Herder's *Alteste Urkunde des Menschen Geschlechts* (1774–1776). The translation is probably by Thomas Holcroft. There was no published English translation of this work at the time and Holcroft refers to the German title. See Lavater, *Essays on Physiognomy*, 3f.

Simultaneously, this symmetry is an image of God's perfection, therefore infinite and ultimately imperceptible.

John Graham has pointed out that Lavater's physiognomy displays three interrelated basic corollaries that anticipate central ideas of Romanticism. First is the claim that all created things are individually unique and distinct from all other things. Second is the idea that every part has the nature and character of the whole. From this follows that each trait of the human body contains the whole character of man, as the smallest works of creation contain the full character of God. Third, and closely related to the previous claims, is the conclusion about the unity of each individual being: particular features are adapted to a particular whole and cannot be transferred to another whole.[11]

This holism constitutes the framework for Lavater's optimism regarding the possibility to "read" a person's character from particular bodily features.[12] His starting-point is the general empiricist claim that external appearances comprise the foundation of all human knowledge. Because all knowledge is by necessity based on external appearances, our knowledge of human character has to be based on the external appearance of the human body.[13] Lavater compares the physiognomist with the farmer who inspects his cornfield or vine stock: the farmer's judgements about the quality of his crops and his expectations concerning his future harvest are all based on the external appearance of the corn or the vine stock.[14] In the same way, the physiognomist makes judgements about human character on the basis of external appearance.

In his analyses of human character, Lavater separates the animal, moral, and intellectual life of a human being. These distinctions are reminiscent particularly of the Platonist tripartite soul composed of a rational (*logistikon*), an acting or "spirited" (*thymoeides*), and an appetitive (*epithymētikon*) part; ideally, they are regulated by the three corresponding virtues: wisdom, courage, and moderation.[15] Lavater emphasises that the animal (or physiological, as he occasionally calls it), moral, and intellectual life unite in one being. These forms or aspects of life do not reside in different parts of the body but rather "coexist in one point" and "form one whole."[16] Nevertheless, this "threefold life of man, so intimately interwoven through his frame, is still capable of being studied in its different appropriate parts."[17] Lavater's

---

[11] Graham, "Lavater's *Physiognomy* in England," 563.

[12] Lavater is cautious, though, and writes: "I do not say that the physiognomist should finally determine by a single sign; I only say it is sometimes possible," cited from Graham, "Lavater's *Physiognomy* in England," 563 n. 8.

[13] Lavater, *Essays on Physiognomy*, 11f.

[14] *Ibid.*, 27f.

[15] Plato, *Republic* IV, 435e–444e, IX, 580d–581a; *Phaedrus* 246a–b, 253c–255b; cf. F. E. Peters, *Greek Philosophical Terms: a Historical Lexicon* (New York: New York University Press, 1967), 170.

[16] Lavater, *Essays on Physiognomy*, 14f.

[17] *Ibid.*, 16.

localisation of appropriate parts echoes the hierarchical localisation Plato presents in the *Timaios*, where he describes how the three parts of the soul are seated in the human body.[18]

According to Lavater, animal life, the lowest and most earthly aspect, "discover[s] itself from the rim of the belly to the organs of generation, which would become its central or focal point,"[19] whereas the middle or moral life is seated in the breast, with the heart as its central point. The most supreme mode of life, the intellect, resides in the head, with the eyes as its centre. Lavater's physiognomist intention is revealed especially by his emphasis on the active and vivid eye as a mark of an active and acute intellect.[20] He presents an additional hierarchical localisation of the three forms of life in the corresponding three parts of the human face. Physiognomical analysis focuses on the countenance and this localisation is therefore of great practical importance to the physiognomist. According to Lavater, the forehead, down to the eyebrows, mirrors the intellect or the understanding; the nose and the cheeks portray the moral life; and the mouth and chin are an image of animal life.[21] The idea of the mouth as a mirror of sensuality occurs especially frequently in Lavater's particular physiognomical descriptions. Interestingly, when Lavater focuses on countenance, he conceives the eyes as the centre of the whole character, combining animal, moral, and intellectual life, and thereby becoming the central characteristic of the unique individual.[22]

The Platonist influence in Lavater's physiognomy is manifest also in his occasional references to the harmony of mind and body. The topic of harmony is central in Plato's *Timaios*, where the ethical harmony of the parts of the human soul and body is an instantiation of the harmonic intervals of the World Soul.[23] Lavater emphasises the value of harmonious relations between the features of the body, especially concerning countenance, and he testifies that there is an evident harmony between physical and moral beauty. But he is also critical of the idea that harmony can be presupposed as a metaphysical idea – a critique that appears to be most directly pointed at Leibniz's perfectionism.[24] According to Lavater, the indicative nature of harmony is something that must be decided on the basis of facts, or "the actual state of things in nature."[25] Consequently, there is a need for physiognomical observation.

---

[18] Plato, *Timaios* 69d–72e.

[19] Lavater, *Essays on Physiognomy*, 16.

[20] *Ibid.*, 23. This physiognomical emphasis on the role of the eye cannot be found in the *Timaios*, though Plato does in other contexts give the eye and sight special attention by describing the intellect or *nous* as the eye of the mind, which is able to perceive eternal truths; see, e.g., Plato, *Republic* VI, 508c–e.

[21] Lavater, *Essays on Physiognomy*, 16f.

[22] *Ibid.*, 17.

[23] Plato, *Timaios* 35b–36b, 69b–d, 90c–d; cf. Peters, *Greek Philosophical Terms*, 176f.

[24] Lavater, *Essays on Physiognomy*, 175ff.

[25] *Ibid.*, 178.

9.2 Physiognomy as Science

The basic question for Lavater's physiognomy is the question of how to combine the individual uniqueness of the particular with the search for universal criteria of interpretation and explanation. He claims that physiognomy is as capable of becoming a science as most other sciences. Only mathematics has a stronger claim to validity.[26] This remark is interesting, because it shows, on the one hand, that Lavater conceives mathematics as the paradigmatic science and, on the other hand, that he defends physiognomy on the grounds that all sciences are lacking compared to mathematics. Thus, physiognomy is no more unscientific than other empirical sciences.

Lavater is – at least rhetorically – self-critical and frequently emphasises that he does not provide a finished and perfected system but only "fragments" of physiognomical studies.[27] He defines scientific knowledge as knowledge that can be explained by principles.[28] In order to become scientific, physiognomy has to meet the challenge of finding its specific principles. The task is difficult, because, as Lavater is acutely aware, individual variation is more apparent than regularities based on kinds.[29] Lavater's emphasis on the unique nature of the individual seems again and again to take the upper hand of his search for universal regularities. The task is further complicated, because very small differences in appearance "scarcely discernible to an inexperienced eye, frequently denote total opposition of character."[30] Lavater's arguments often give the impression that he wants to save the rule by allowing an infinite amount of exceptions. This strategy markedly defeats the claim that physiognomy is based on universal principles.

As we saw above in relation to the case of harmony, Lavater claims that metaphysical structure must not be postulated but disclosed by empirical observation. He believes in the method of induction, but his inductionism is based on the philosophical assumption that there are observable regularities in nature. Lavater claims that the methodology of physiognomy is based on the fact that there is a natural and universal tendency, shared by humans and animals, to judge things by their appearance. This universal tendency is not as such a developed scientific physiognomy but "the confused feeling of the agreement between the internal character and the external form."[31] As evidence for this tendency, Lavater points at "the number of physiognomical terms to be found, in all languages, and among all nations."[32] G.P. Brooks and R.W. Johnson have pointed out that this kind of *consensus genitum*

---

[26] *Ibid.*, 67.
[27] *Ibid.*, 18.
[28] *Ibid.*, 68.
[29] *Ibid.*, 22.
[30] *Ibid.*, 99.
[31] *Ibid.*, 58.
[32] *Ibid.*, 58f.

argument was typical of the period and adopted, among others, by Thomas Reid and his successors in the Scottish School.[33]

Lavater's account of physiognomical judgment can also be illuminated by a comparison with the notion of reflective judgments, which Kant develops in *Critique of the Power of Judgment* (1790). Kant's prime interest in this third critique is to study the possibility of judgments of taste. These judgments are characteristically subjective and particular, and they cannot be verified, but they nevertheless lay claim to the agreement of other subjects of experience.[34] According to Kant, "[t]he power of judgment in general is the faculty for thinking of the particular as contained under the universal."[35] He separates between *determinant* judgment, where the particular is subsumed under a pre-given universal rule, principle, or law and *reflective* judgment, where "only the particular is given, for which the universal is to be found" (Ak 5:179). Kant explicitly distinguishes the principle of reflective judgment from, on the one hand, the universal laws of nature, which are grounded in the understanding, and, on the other hand, the universal practical law, which determines the moral will (Ak 5:180f). The principle of a reflective judgment is universal in a weaker sense than the laws of nature and morality: it is valid for only one particular judgment. It is universal, because it "rightly makes claim to the assent of everyone else, even though this judgment is empirical and is an individual judgment" (Ak 5:191). The rightly laid claim to agreement is, according to Kant, based on the transcendental principle of the formal finality of nature. This finality cannot be captured by any concept or determined a priori but can be grasped only in a judgment of taste, to which it appears as beauty (Ak 5:181ff).

There is an interesting similarity between Lavater's struggle to combine the unique nature of particular physiognomies with some kind of universal structure and Kant's study of reflective judgment. Even more so, a comparison between Lavater and Kant illuminates in what sense physiognomy is pseudo-scientific. According to Kant, taste constitutes the paradigmatic case of reflective judgment, and its realm is strictly distinguished from, respectively, the realms of science and morality. He explicitly states that "the judgment of taste is not a judgment of cognition" (§ 38, Ak 5:290) and neither is it concerned with the good (§ 4, Ak 5:207ff). From this perspective, Lavater's physiognomy is pseudo-scientific, because it conflates scientific principles with principles belonging to the realm of aesthetics.

The case of mathematics illuminates the issue. As we saw above, Lavater wants to combine his conception of mathematics as the highest science with the idea

---

[33] Brooks and Johnson, "Lavater's *Essays on Physiognomy*," 9. Brooks and Johnson examine Lavater's method of induction in relation to the scientific ideals of his contemporaries as well as in relation to later psychological theory. For a discussion of the strengths and weaknesses in Reid's inductivism, see Riku Juti's article in this volume.

[34] Cf. Eva Schaper, "Taste, Sublimity, and Genius: The Aesthetics of Nature and Art," in Paul Guyer (ed.), *The Cambridge Companion to Kant* (Cambridge: Cambridge University Press, 1992), 371.

[35] Immanuel Kant, *Critique of the Power of Judgment*, trans. by Paul Guyer and Eric Matthews (Cambridge: Cambridge University Press, 2000), 66; Akademie edition vol. 5:179. Further references are given to the Akademie edition; all quotations are from Guyer's translation.

that all particular physiognomies are unique but still capable of being made the objects of science. Kant emphasises the profound difference between the realms of mathematics and aesthetics. Mathematics is based on concepts, and when drawing a geometrical figure the imagination is determined by the understanding. In aesthetic judgment, by contrast, the imagination is free and only accidentally coincides with the understanding when we judge something to be beautiful (Ak 5:240f). From this accidental nature follows, among other things, that we cannot beforehand determine that something is going to be judged to be beautiful (Ak 5:191). Kant markedly draws the conclusion that geometrical properties, such as symmetry or perfection, are not beautiful and that it is an often-conducted mistake to claim so. True beauty is always to some extent irregular and non-symmetrical (Ak 5:240ff). From Kant's perspective, the objects of particular reflective judgment are unique, exactly because they are not determined by any principle, law, or concept. From this viewpoint, Lavater's attempt to combine scientific principles and unique physiognomies becomes a kind of category mistake.

9.3 The Corrective Impact of Errors

Lavater readily admits that the physiognomist makes errors, but he claims that these errors do not imply that physiognomy as such is untrue. Errors mean that the discernment of the physiognomist errs, not that countenance lies about the character of a person. Lavater claims that "to conclude there is no such science as physiognomy, because physiognomists err, is the same thing as to conclude there is no reason, because there is much false reasoning."[36] Errors provide the physiognomist with a possibility to correct and extend his or her knowledge. Lavater writes, "I have been, and continue daily to be, mistaken in my judgement; but [. . .] these errors are the natural, and most certain, means of correcting, confirming, and extending my knowledge."[37] This emphasis on the positive impact of errors has been interpreted as a rudimentary principle of falsification,[38] but I think it is more appropriately understood as an aspect of Lavater's emphasis on the self-critical and essentially moral task of becoming a physiognomist.

Lavater's remarks on the task of becoming a physiognomist and on the nature of physiognomical perception are often philosophically and psychologically more interesting than his physiognomical analyses of countenances. He points out, for example, that we only seldom avoid viewing things through the medium of our own inclinations and aversions:

> Obscure recollections of pleasure or displeasure, which this or that countenance have by various incidents impressed upon his mind, impressions left on his memory, by some object of love or hatred – How easily, nay, necessarily, must these influence his judgement! Hence,

---

[36] Lavater, *Essays on Physiognomy*, 128.

[37] *Ibid.*, 6f.

[38] Brooks and Johnson point in this direction. See "Lavater's *Essays on Physiognomy*," 8.

how many difficulties must arise to physiognomy, so long as physiognomy shall continue to be the study of men and not angels![39]

Men cannot become angels, and therefore the practice of self-knowledge is the only possibility of avoiding and overcoming prejudiced judgements. Self-knowledge, according to Lavater, "that most essential yet most difficult of all knowledge, to the physiognomist, ought to be possessed by him in all possible perfection. In proportion only as he knows himself will he be enabled to know others."[40] Lavater's emphasis on self-knowledge highlights the fact that the task of the physiognomist is a moral as much as a scientific task. It also shows that when it becomes difficult to articulate the physiognomical *method* in any detail, Lavater shifts emphasis and focuses on the character and abilities of the physiognomist.

Michael Shortland has argued that the ingenuity of Lavater's physiognomy is best revealed when it is seen as an attempt to articulate a method of physiognomical *perception*.[41] Physiognomical perception is distinguished from physiognomical sensation, which is the universal capacity to judge content on the basis of form and which is shared by all creatures that have eyes. Physiognomical perception allows the perceiver to *think* rather than feel physiognomically and this capacity can be developed only by humans.[42] To become an excellent physiognomist requires inborn talent, according to Lavater, but above all one has to practice one's talent by self-critical use.[43] The *Essays on Physiognomy* are principally intended to enlarge the reader's experience. The number of physiognomies and characters anyone meets during a lifetime is limited and therefore it is important to use art, fiction – and physiognomical handbooks – to extend and vary the experience provided by daily life.

9.4 The Science and Art of Genius

Lavater seems to think that physiognomy is primarily learned through informed experience rather than through the study of principles. Physiognomical knowledge is characteristically intuitive rather than deductive. Interestingly, Lavater claims that physiognomical intuition is to a large extent acquired. He explicitly refers to an "intuition of experience".[44] Lavater's concept of intuition seems to combine a rationalist conception of non-deductive understanding or mathematical intuition with an early Romanticist conception of artistic creative intuition. This hybrid of intuition provides a key to Lavater's understanding of physiognomy as well as to its most profound problems. He oscillates between science understood as knowledge that

---

[39] Lavater, *Essays on Physiognomy*, 104.

[40] *Ibid.*, 123.

[41] Michael Shortland, "Skin Deep: Barthes, Lavater and the Legible Body," *Economy and Society* 14:3 (1985), 285f.

[42] Lavater, *Essays on Physiognomy*, 106.

[43] *Ibid.*, 92f.

[44] *Ibid.*, 113. Here Holcroft gives the German term "der Geübte intuitif".

can be explained by principles and science understood as a branch of creative art. He claims that, like all sciences, physiognomy "may, to a certain extent, be reduced to rule and acquire an appropriate character, by which it may be thought." But as in "every other science, so, in this, much must be left to sensibility and genius."[45] Interestingly, Lavater emphasises that scientific discoveries and philosophical inventions are not achieved by following rules but rather by using creative capacities. The inventive scientist or philosopher will "make numerous remarks, that are not reducible to rule."[46]

Lavater goes far in his emphasis on the creative aspect of science and overtly fuses science and art. The fusion seems to be based on the fact that Lavater does not separate the truth of the sciences from the truth of the creative arts. He uses the difference between the Renaissance artists Albert Durer and Raphael as an example:

> Albert Durer surveyed and measured men: Raphael measured men still more feelingly than Albert Durer. The former drew with truth, according to rule; the latter followed his imagination; yet was nature often depicted by him with not less exactness.[47]

When physiognomy is a science based on rules it measures like Durer, according to Lavater, and when it is based on the perception of a genius it depicts the truth like Raphael. Occasionally, Lavater seems to think physiognomy has to revert to the perception of geniuses, because the science is new and has not yet developed an appropriate set of principles,[48] but more often he appears to claim that the ultimate truth of physiognomy, like the ultimate truth of painting, can be grasped only by a genius and not even in principle reduced to principles.

It is illuminative to compare Lavater's conception of genius with Kant's remarks on the topic. According to the third *Critique*, "*Genius* is the talent (natural gift) that gives the rule to art" (§ 46, Ak 5:307). Talent belongs to nature and genius and is thus the innate mental aptitude through which nature gives the rule to art. The products of genius are characterised by originality, and they are, thus, by definition not based on pre-given rules. They are distinguished from nonsense by their exemplary nature: they are models, which set the standard of estimating. This standard is a kind of rule but a rule that cannot be set down in a formula or serve as a precept, because then judgments of the beautiful would be determinable according to concepts (§§ 46–47, Ak 5:307ff). Whereas taste is the capacity to estimate beautiful things, genius is the capacity to produce such things (§ 48, Ak 5:311). Kant emphasises that genius cannot be learned, because learning is based on repetition, while genius is always novel, but this does not mean that the learning of rules has no role in fine art. He writes that "[g]enius can only provide rich *material* for products of art; its elaboration and *form* require a talent that has been academically trained, in order to make a use of it that can stand up to the power of judgment" (§ 47, Ak 5:310).

---

[45] *Ibid.*, 67.
[46] *Ibid.*, 71.
[47] *Ibid.*, 72.
[48] *Ibid.*, 67, 69.

Kant's conception of the genius as giving the rule by producing exemplary models illuminates the role Lavater seems to attribute to the genius of physiognomical knowledge. But again, it is crucial to keep in mind that, contrary to Lavater, Kant makes a strict distinction between the realms of science and aesthetics. He distinctly withholds that genius belongs only to art, not to science. This is so, exactly because science is based on rules, which can be learned and followed once they have been discovered: "Newton could make all the steps that he had to take, from the first elements of geometry to his great and profound discoveries, entirely intuitive not only to himself but also to everyone else" (§ 47, Ak 5:309). Kant claims that the difference between a scientific inventor, such as Newton, and an ordinary practitioner of science is a difference of degree only, whereas the difference between a genius and someone who creates according to pre-given rules is a qualitative difference. His conclusion thus excludes creativity from science and shows that he has to pay a rather high price to maintain a strict distinction between the realms of science and aesthetics. Lavater's blurring of the distinction recognises the creativity of scientific discovery, but as we have seen it gives rise to other problems.

In this context, one should also keep in mind the theological aspect of Lavater's thought. The symmetry between appearance and character is created by God, who has made humans images of his own perfection, and ultimately this symmetry escapes the complete grasp of finite humans, who always remain prone to error. The genius is someone who is not chained by human finitude and is able to approach the limit value of God's perfection. From the intuitive nature of physiognomical knowledge follows the difficulty of communication. Using Kantian terminology, one can say that this knowledge cannot be communicated by concepts. Lavater is aware of the problem of communicability and claims that the art of drawing has to step in where words fail:

> Drawing is the first, most natural, and most unequivocal language of physiognomy; the best aid of the imagination, the only means of preserving and communicating numberless peculiarities, shades, and expressions, which are not by words, or any other mode to be described.[49]

Therefore, the art of drawing is indispensable to the physiognomist, who will not otherwise be able "to make, much less to retain, or communicate, innumerable observations."[50] These physiognomical drawings have a role that resembles the exemplary models created by the Kantian genius: they give rules that cannot be set down in formula and can therefore be expressed only by the drawings themselves. But not even a vivid imagination and developed skill of drawing is always enough, according to Lavater. There are physiognomical perceptions that cannot be retained by the imagination and are "rather felt than seen."[51] Examples of such ungraspable aspects are "the look of love," "the soft benignant vibration of the benevolent eye"

---

[49] Lavater, *Essays on Physiognomy*, 122, cf. 69.
[50] *Ibid.*, 122.
[51] *Ibid.*, 100.

and "the internal strong efforts of the mind."[52] Lavater maintains that these can be perceived even though they can neither be described in any detail nor expressed even by the most skilled artist.

9.5 The Accidental and the Necessary

As mentioned, Lavater distinguishes physiognomy and pathognomy. Whereas physiognomy studies stable features of bodily form, pathognomy studies the expression of passions and bodily movements. Lavater writes that "[p]hysiognomy, opposed to pathognomy, is the knowledge of the signs of the powers and inclinations of men. Pathognomy is the knowledge of the signs of the passions. Physiognomy, therefore, teaches the knowledge of character at rest; and pathognomy of character in motion."[53] From our contemporary perspective, pathognomy as a study of bodily expression would seem to provide a more promising perspective on human character, but for Lavater it is the other way around. The topic of deception and hypocrisy was frequently used in the fiction of the time and the interpretation of bodily expression was seen as problematic because people are able to manipulate their expressions and deceive.[54] Physiognomy is more scientific than pathognomy, because a study of a body at rest and of bodily constitution excludes the possibility of deception. Wollstonecraft will take this idea to its morbid extreme in her story "The Cave of Fancy", where the main character studies corpses, which are seen as especially revealing and truthful.[55]

Still, the distinction between physiognomy and pathognomy does not seem to be absolute. Lavater mentions that these two sciences are inseparable to "the friend of truth," and he motivates his own focus on physiognomy by the fact that it has been less studied than pathognomy.[56] It is also important to note that contrary to what Wollstonecraft claims in her review of Shaw's abridgement, Lavater's physiognomy is not solely focusing on "the original formation of the bones."[57] The mouth plays an important role in many of Lavater's reflections on particular countenances and here the formation is only partly directed by bone-structure. The distinction between physiognomy and pathognomy is, rather, concerned with the distinction between the countenance at rest and the countenance affected by passions. But Lavater does not

---

[52] *Ibid.*

[53] *Ibid.*, 20.

[54] *Ibid.*, 21. For discussions of the relation between Lavater's writings and eighteenth-century fiction, see Shortland, "Skin Deep," 279–86; and Rivers, *Face Value*.

[55] It is interesting to note that Wollstonecraft's daughter Mary Shelley took up the topic of corpses and bodily parts in her novel *Frankenstein* (1818). I thank Sara Heinämaa for pointing out this connection.

[56] Lavater, *Essays on Physiognomy*, 21. This is noted in the review of Shaw's abridgment, where Wollstonecraft quotes this claim from Holcroft's translation; see Wollstonecraft, "Contributions to the *Analytical Review*," 450.

[57] *Ibid*.

seem to regard even this distinction as absolute. One of his prime examples of the correspondence between appearance and character concerns the correlation between the eye and the mind. He writes:

> After repeated observation that an active and vivid eye and an active and acute wit are frequently found in the same person, shall it be supposed that there is no relation between the active eye and the active mind? Is this the effect of accident? – Of accident! – Ought it not rather to be considered as sympathy, an interchangeable and instantaneous effect, when we perceive that, at the very moment the understanding is most acute and penetrating and the wit the most lively, the motion and fire of the eye undergo, at that moment, the most visible change?
>
> Shall the open, friendly, and unsuspecting eye, and the open, friendly, and unsuspecting heart, be united in a thousand instances, and shall we say the one is not the cause, the other the effect?[58]

This focus on the active eye does not primarily concern the countenance at rest, but still Lavater considers it as an example of physiognomy rather than pathognomy. The reason for this seems to be that the activity of benevolence of one's eye is something one cannot feign. In a chapter dedicated to the problem of dissimulation, Lavater explicitly claims that the lustre, as the colour, of one's eyes is something one cannot alter at one's pleasure.[59]

The distinction between physiognomy and pathognomy can be seen as resting on a more basic distinction, which Lavater makes between the necessary and the accidental. He believes that there are necessary causes in nature and this causality grants, among many other things, the correlation between character and appearance. A certain physiognomy is the necessary effect of a certain character. Passions are also necessary effects of certain causes, but these causes do not have a necessary relation to character, and therefore passions are in this respect merely accidental.[60]

In a chapter on the question of freedom and necessity, Lavater uses a telling example to describe the range of human freedom. He writes: "man is as free as the bird in the cage; he has a determinate space for action and sensation, beyond which he cannot pass. As each man has a particular circumference of body, so has he likewise a certain sphere of action."[61] The human cage is constituted by one's character, which is given and expressed in one's physiognomy. This claim is explicitly directed against the Enlightenment philosopher Claude-Adrien Helvétius, one of whose "unpardonable sins" was to assign "to education the sole power of forming, or deforming, the mind."[62] Still, Lavater does not deny that there is an

---

[58] Lavater, *Essays on Physiognomy*, 23f.

[59] *Ibid.*, 154.

[60] *Ibid.*, 142. Lavater seems to think that character is the primary cause and appearance its effect, but he is not very explicit about the direction of the causality and mostly comments only on the necessary correlation between character and appearance.

[61] *Ibid.*, 166.

[62] *Ibid*.

impact of education, and he discusses at some length how improvement as well as degradation changes the character and thereby also the physiognomy of a person.[63]

Human freedom and virtue consist in this possibility of improvement or degradation, but this possibility is always restricted by what is given. Lavater argues against the Enlightenment conception of universal man, who has equal opportunities, and defends the idea that every person is a unique individual, stating that "[e]ach individual can but what he can, is but what he is. He may arrive at, but cannot exceed, a certain degree of perfection, which scourging, even to death itself, cannot make him surpass. Each man must give his own standard."[64] Lavater defends a conception of morality that is based on necessity and particularity, rather than on the Enlightenment principles of freedom and universality. He makes an interesting comparison with musical instruments: each instrument is capable of an infinite variety of sounds: it can be played well or less well, and its sound can be perfected. But a violin can never become a flute.[65] This view of original constitution is challenged by Wollstonecraft's emphasis on the role of education. I will now conclude by making some remarks on her critical interpretation of Lavater's physiognomy.

9.6 Bodily Constitution, Equality, and Education

In her treatment of the nature of virtue, Wollstonecraft refers repeatedly to the close connection between bodily constitution and character. Several commentators have related this aspect of her moral thought to Lavater's physiognomy. The recent biographer Janet Todd claims that Wollstonecraft's work on the translation of Lavater's physiognomy "accelerated her sense of the mind and body's interdependence."[66] I think it might be exaggerated to detect a direct influence here, but I do think that Wollstonecraft's interest in the connection between bodily constitution and character illuminates her fascination with Lavater's theory.

It is significant, though, that Wollstonecraft's only direct application of Lavater's ideas is found in an unfinished and posthumously published work of fiction "The

---

[63] Most of the chapter "On the Harmony of Moral and Corporeal Beauty" is dedicated to this question, see *ibid.*, 175–203. It is also important to note that the given character is in itself neither vicious nor virtuous: children are born innocent. Lavater emphasises that in this moral sense nobody has either good or bad dispositions; see *ibid.*, 215.

[64] *Ibid.*, 166f.

[65] *Ibid.*, 169.

[66] Janet Todd, *Mary Wollstonecraft: A Revolutionary Life* (London: Weidenfeld & Nicolson, 2000), 134. See also Sapiro, *A Vindication of Political Virtue*, 20–21, 68. There is also a personal connection between Wollstonecraft and Lavater, which has fascinated biographical commentators; see Virginia Sapiro, *A Vindication of Political Virtue: the Political Theory of Mary Wollstonecraft* (Chicago: University of Chicago Press), 20–21. At the end of the 1780s, when she translated Lavater's work, Wollstonecraft was in love with the Swiss painter Henry Fuseli, who admired and knew Lavater well. I do not think this biographical fact should be used in order to explain Wollstonecraft's interest in Lavater's theory. She often seems to have chosen her amours on the basis of her philosophical ideas rather than the other way around.

Cave of Fancy" probably written at the end of the 1780s, simultaneously or shortly after her work on the translation of Lavater's work. This story presents humans at the mercy of nature: a shipwreck kills all the passengers but a young girl, who is saved and educated by a sage, Sagestus, who is aided by benevolent spirits of nature. The sage studies the features of the victims in order to estimate their character. He is especially interested in studying the corpse of the mother of the child he has saved. Wollstonecraft writes:

> On observing her more closely, [Sagestus] discovered that her natural delicacy had been increased by an improper education, to a degree that took away all vigour from her faculties. And its baneful influence had had such an effect on her mind, that few traces of the exertions of it appeared on her face, though the fine finish of her features, and particularly the form of the forehead, convinced the sage that her understanding might have risen considerably above mediocrity, had the wheels ever been put in motion;[67]

This passage interestingly emphasises the visible traces left by a bad education and introduces the topic of women's degradation by education, which Wollstonecraft develops in *A Vindication of the Rights of Woman* (1792). Concurrently, the impact of education is combined with the idea that there are inborn individual differences in ability, which can still be seen behind the acquired features. Somehow surprisingly, the sage's interest in the mother of his protégée turns out not to be directed by the idea that the girl has inherited her mother's abilities. Instead, he draws the conclusion "that the orphan was not very unfortunate in having lost such a mother,"[68] adopts the girl, and gives her the name Sagesta after himself.

By attributing the perceptive capacity to somebody more than human, somebody who "overleaped the boundary prescribed to human knowledge,"[69] the fictional character of Sagestus allows Wollstonecraft to solve Lavater's problem of the limits that human finitude poses on physiognomical perception. Wollstonecraft repeatedly tells the reader that Sagestus' physiognomical knowledge is based on the fact that he is allowed "to see a great way farther than common mortals."[70] Due to his supra-human character, he is "perfectly acquainted with the construction of the human body"[71] and able to abandon conjecture in favour of distinct ideas.[72]

In the "Cave of Fancy", Wollstonecraft takes a double distance to Lavater's physiognomy: first by staging the theory in a fictional context and second by attributing physiognomical knowledge to someone more than human. When her disparaged remarks on physiognomy and related issues are taken as a whole, Wollstonecraft seems to agree with Lavater's basic idea that there is a close connection between bodily constitution, appearance, and character. But she puts more emphasis on the idea that both constitution and character can be changed by education. Most

---

[67] Mary Wollstonecraft, "Cave of Fancy," in *The Works of Mary Wollstonecraft*, vol. 4, eds. Janet Todd and Marilyn Butler (London: William Pickering, 1989), 197.

[68] *Ibid.*, 198.

[69] *Ibid.*, 191.

[70] *Ibid.*

[71] *Ibid.*, 194.

[72] *Ibid.*, 197f.

importantly, she seems to think that *if* there is an exact correspondence between character and features of countenance, this correspondence transcends finite human knowledge and can be known only by someone who has perfect knowledge of nature and the aims of its creator. Humans who try to interpret exact physiognomies easily get trapped in ridiculous conclusions – as Wollstonecraft claims was the case with Lavater himself in his view on powerful potential benefactors.[73]

The formative capacity of education plays a more significant role in Wollstonecraft's than in Lavater's thought. In this respect she is closer to Helvétius and the educational optimism of the Enlightenment. In the *Rights of Woman*, the question of education becomes the most important feminist issue. The fact that women are inferiorly educated explains that their achievements do not always reach the level of men's achievements. Wollstonecraft does think that a person's character is to a great extent formed by education, understood in a broad sense, including all kinds of external impressions. In the *Rights of Woman* she makes the point by comparing women and soldiers:

> As a proof that education gives this appearance of weakness to females, we may instance the example of military men, who are, like them, sent into the world before their minds have been stored with knowledge or fortified by principles. The consequences are similar; [. . .] Soldiers, as well as women, practise the minor virtues with punctilious politeness. Where is then the sexual difference, when education has been the same?[74]

Wollstonecraft emphasises the role of education also in relation to the connection between character and constitution. Her conception of bodily constitution is significantly less static than Lavater's. Her interest in the mind–body connection is most evident when she emphasises the importance of physical exercise for children of both sexes. Wollstonecraft writes:

> Let fancy now present a woman with a tolerable understanding, [. . .] whose constitution, strengthened by exercise, has allowed her body to acquire its full vigour; her mind, at the same time, gradually expanding itself to comprehend the moral duties of life, and in what human virtue and dignity consists.[75]

The focus on exercise implies that bodily constitution is something that can be changed – educated in its own right – and this focus is quite different from Lavater's static view of the body as a cage.

Given Wollstonecraft's belief in the power of education, her enthusiasm for Lavater's thought does seem odd. It is evident that she did at some point distance

---

[73] In *Letters Written during a Short Residence in Sweden, Norway, and Denmark* (1796), the last work Wollstonecraft published during her lifetime, she mentions Lavater in connection with the Danish Foreign Minister Andreas Peter Bernstorff, whom she had met and who had two years earlier invited Lavater to Denmark. The comment ridicules Lavater, who "has a knack at seeing a great character in the countenances of men in exalted stations, who have noticed him, or his works." See *The Works of Mary Wollstonecraft*, vol. 6, eds. Janet Todd and Marilyn Butler (London: William Pickering, 1989), 331.

[74] Mary Wollstonecraft, *A Vindication of the Rights of Woman*, in *The Works of Mary Wollstonecraft*, vol. 5, eds. Janet Todd and Marilyn Butler (London: William Pickering, 1989), 92.

[75] *Ibid.*, 119.

herself from the conservative and speculative implications of Lavater's theory, and it is possible that this distancing concurred with the radicalizing effect the French revolution – and especially its English reception – had on her thought.[76] There is no reference to Lavater and no application of physiognomical ideas in the *Rights of Woman*, even though the book was published in the same year as the review of Shaw's abridgment.

However, it is important to note that Wollstonecraft is not afraid of adopting ideas from philosophers she also criticized, such as Jean-Jacques Rousseau, or to use elements of thought from philosophers with radically opposite views. In her moral philosophy, she attempts to combine aspects of the Reverend and mathematician Richard Price's rationalist, universal, and explicitly Platonist moral realism with Rousseau's emphasis on particularity and the essential role of passions and imagination in human morality.[77] In the case of Rousseau, she agrees on most parts of his educational program even though she is untiring in her critique of his conclusions concerning the difference between male and female virtue. From this perspective, it is not so surprising that she is both inspired by and critical towards Lavater's physiognomy.

Wollstonecraft's claim that sexual difference is based on education does not mean that there are no differences between individuals. It is not always entirely clear whether she claims that there is no inborn sexual difference relevant to a person's character, or rather that education is able to compensate differences. According to her comparison between women and soldiers, sexual difference is claimed to be created by education (or rather the lack of proper education).

But Wollstonecraft does also write: "I have already granted, that from the constitution of their bodies, men seem to be designed by Providence to attain a greater degree of virtue."[78] She emphasises that even though there are differences between individuals, and men often seem to be given a better starting point due to their stronger bodies, this does not mean that virtue has different standards for men and women. Wollstonecraft's conception of universal equality is strongly tied to her

---

[76] In her comment on Lavater in *Short Residence*, Wollstonecraft remarks that Lavater and Bernstorff agreed in their "sentiments relative to the French revolution" (331). Bernstorff is best known for his politics of neutrality, which involved not taking a stance on either the French revolution or the war between France and England in order to protect Danish interests in trade. Wollstonecraft's comment seems to imply that they shared a conservative view on this issue – but from the point of view of the revolutionary, neutrality might well appear conservative.

[77] I have examined Rousseau's influence on Wollstonecraft's thought in my "Mary Wollstonecraft on Love and Friendship," in Åsa Carlson (ed.), *Philosophical Aspects on Emotions* (Stockholm: Thales, 2005), 119–39; and "Catharine Macaulay and Mary Wollstonecraft on the Will," in Jacqueline Broad and Karen Green (eds.), *Virtue, Liberty, and Toleration: Political Ideas of European Women, 1400–1800* (Dordrecht: Springer, 2007), 149–69. I have studied Price's influence on Wollstonecraft's critique of Rousseau in my "Mary Wollstonecraft on the Foundation and Acquisition of Virtue," paper presented at the conference *Women, Metaphysics and Enlightenment, 1660–1789*, 24th March 2006, Institute of Philosophy, London.

[78] Wollstonecraft, *Rights of Woman*, 95.

view that "virtue has only one eternal standard"[79] that everyone should be given equal opportunity to approach. Her claim about one eternal standard is the opposite of Lavater's claim that "[e]ach man must give his own standard,"[80] but it is important to note that Wollstonecraft's universalism does attempt to take into account the existence of inborn as well as acquired differences between individuals.

Here, Wollstonecraft expresses an implicit critique of Price's universalism and struggles to combine the universal nature of virtue with the unique nature of moral situations. Wollstonecraft emphasises that women cannot become truly virtuous by obeying rules and must learn to practice their reason independently, because "in every circumstance of life there is a kind of individuality, which requires an exertion of judgement to modify general rules."[81] She attempts to develop an understanding of moral life that takes into account the demand for universal standards and equality as well as existing unique differences in character and situation. Lavater's emphasis on the unique individuality of bodily constitution and character does correspond to Wollstonecraft's recognition of particularity.

---

[79] *Ibid.* Here Richard Price's influence on Wollstonecraft's thought is especially evident, see Richard Price, *A Review of the Principal Questions in Morals*, 3rd ed. (London: T. Cadell in the Strand, 1787), 277–78. Based on information from the *English Short Title Catalogue. Eighteenth Century Collections Online*. Gale Group. http://galenet.galegroup.com/servlet/ECCO

[80] Lavater, *Essays on Physiognomy*, 167.

[81] Wollstonecraft, *Rights of Woman*, 249.

Chapter 10
Toward the Rebirth of Aristotelian Psychology: Trendelenburg and Brentano

Edoardo Fugali

Abstract The chapter studies the concepts of the self, the soul and the subject as they were developed around the first half of the nineteenth century in German philosophy presiding over the birth of psychology as a science. The topic is addressed by examining particularly the leading roles that Friedrich Adolf Trendelenburg and Franz Brentano played in this development. Both thinkers worked out an original conception of the soul through recourse to Aristotle's theories of the soul, combining them with insight stemming from the modern tradition of the philosophy of subjectivity, particularly Kantianism. The first part of the chapter explicates Friedrich Trendelenburg's argument, that psychology constitutes an independent discipline, and show how his arguments contributed to the general discussion about the status of psychology. The second part consists of an explication of Franz Brentano's reinterpretation of the Aristotelian tradition and provides a critical comparison between his position and that of Trendelenburg. The main argument of the chapter is that Trendelenburg had an important mediating role in the post-Aristotelian tradition, which developed further and culminated in idealistic theories of subjectivity and self-consciousness.

Keywords Descriptive psychology · Aristotle's *De anima* · sense perception · imagination · Brentano · Trendelenburg

There are good reasons to argue that both Friedrich Trendelenburg and Franz Brentano worked out an original conception of the soul through recourse to Aristotle's psychological theories of the soul and its constitution, combining them with insights stemming from the modern tradition of the philosophy of subjectivity, particularly Kantianism. The importance of Aristotelian concepts for these thinkers will be discussed in this chapter both with regard to topics and methods, and the conceptual question of the soul will be related to a larger problem concerning the epistemological status of psychology and its demarcation from and dependency on philosophy. The latter issue can be summed up by asking: Can psychology be

E. Fugali
Department of Philosophy, University of Trieste

established as an autonomous science, exclusively relying on empirical methods and on the evidence given by the physiology of the sense organs, independently of the concept of the soul? Or must psychology rather restore its relationship with philosophy in order to liberate itself from the theoretical impasses to which such methodological restrictions seem to condemn it?

Within this systematic and historic framework, I will first explicate Trendelenburg's argument that psychology constitutes an independent discipline, and show how his arguments contributed to the general discussion about the status of psychology and functioned as a mediating position in this crucial period. It must be emphasized, above all, that Trendelenburg's approach has two distinctive features: he argues that the relation between psychology and metaphysics is organic, and he strives to rescue the concept of the soul. Both these features have Aristotelian roots and both work against the Cartesian tradition. The main aim of the first part is to explicate the specific character of Trendelenburg's approach from a historical and critical point of view and to disclose his contribution to the creation of a viable program for psychological research in general and the study of cognitive processes in particular. His approach differs greatly from the current paradigm of experimental psychology, and it gave birth to Brentano's psychology of act and phenomenological psychology.

Trendelenburg's contribution is that he argued, on the basis of a critique of Hegel's dialectic, that logic and its categories have an empirical and factual origin in the structures of language. He investigated these structures from the standpoint of a theory of knowledge which was directly inspired by the comparative folk psychology (*Völkerpsychologie*) of Moritz Lazarus and Heymann Steinthal. These new approaches to the epistemological problem of logic diffused broadly into German philosophy during the second half of the nineteenth century and effected a clear change of perspective: According to these approaches, logic was not to be restricted to an account of an inventory of norms and formal rules, but should be understood as a general propaedeutics and a universal methodology for all specific disciplines and their applications and as such should also include a genetic investigation of the conditions that preside over its constitution. In Trendelenburg's opinion, studying logic from this new perspective is equivalent to studying the psychological structure of experience and the conditions that make it possible. In his epistemological viewpoint, psychology is a methodological study of the conditions that constitute the ground of all subjective knowledge.

The second part of this chapter consists of an explication of Franz Brentano's contribution to the Aristotelian tradition and of a critical comparison between his position and that of Trendelenburg. There is a historical link between these two thinkers: Brentano followed Trendelenburg's lectures on psychology in Berlin in 1858 and 1859. Trendelenburg's influence on Brentano's thought is clarified by an excursion on Brentano's unpublished manuscripts, and especially his lecture notes on Trendelenburg's philosophy. So the main argument of this chapter is that Trendelenburg had an important mediating role in the post-Aristotelian tradition, which developed further and culminated in idealistic theories of subjectivity and self-consciousness.

10.1 Trendelenburg and the Renewal of Aristotle's Psychology

Trendelenburg struggles to introduce and ground an empirical concept of the subject, a concept that would be part of a broader system with a specific metaphysical character. More precisely, the themes of the soul and human subjectivity are brought back to and grounded on an organic concept of the world, which emphasizes and highlights the relations of interaction between the soul and the different levels in which reality is structured.

Trendelenburg's psychological theories can be found in his manuscripts, collected in an extensive series which also contains his notes for the psychology lectures he gave in Berlin between 1840 and 1870.[1] Before focusing on these sources, I want to point out that Trendelenburg's arguments for psychology as an autonomous discipline are very different from and work against the empiricist notion of psychology as an experimental science, a notion that will be established in Germany from 1879 onwards. Trendelenburg's proposal belongs under the heading of "philosophical psychology", and it provides the historical background for Brentano's and Wilhelm Dilthey's descriptive empirical psychologies. More specifically, Trendelenburg's theory of constructive movement anticipates Brentano's reflections on intentional reference and in-existence, and his account of the relation between an individual consciousness and the structures of objective spirit connects him to Dilthey's philosophy and to Wilhelm Wundt's reflections on comparative folk psychology (*Völkerpsychologie*).

Both these aspects (the intentional character of constructive movement and the relationship between individual and objective spirit) distinguish Trendelenburg's concept of empirical psychology from empirical psychology understood as theorization based on controlled experiments and systematic observations. The difference between Trendelenburg's approach and experimential-theoretical psychology cannot be explained simply by the distinction between internal and external, for in a certain sense, both focused on and privileged the internal life of the subject. But as experimental psychology connected introspection with physiological experimentation, it reduced inner life to mere sensory and perceptual processes and, in conformity with the methodological standards of the time, assumed that only these elementary processes, and not the higher functions of the mind, are subject to experimental research.

In contrast to this, Trendelenburg's reflections on psychology as a science and his definition of the object of this science are highly interesting. According to him, the human soul is the prime research object of psychology. The soul holds an intermediate position between nature and the self-objectifications of the spirit, that is, the productions in which the spirit or the mind objectifies itself. The essence of the soul unfolds itself in both these phenomenal fields: nature is structured and formed by sense contribution stemming from the productive activities and force of

---

[1] Friedrich Adolf Trendelenburg, unpublished lectures on psychology, Ms. B 4, Trendelenburg-Nachlass, Staatsbibliothek zu Berlin, preußischer Kulturbesitz.

the spirit. This double condition is reflected in the ambiguous position of psychology as science; the double task of psychology is to understand nature in its intentionality and the spiritual world in its motivating power function. The soul has a double relationality; on the one hand, it relates to the natural world and, on the one hand, it relates to the spiritual world. Its actions manifest this double relation, and this double relation, in its turn, manifests its essence. The relation is not any external and contingent datum, but is immanent and reflective, since the soul relates to itself and manifests its essence through this relation: it presents itself as a striving that leads nature to completion and that allows the emergence of the meaning that the spirit constitutes for the world.

Psychology has a privileged position and function among the sciences. Each science represents the whole of being in its own right, but in Trendelenburg's system of sciences, psychology functions as an intermediative science, which mediates between the sciences of nature (above all physics) and ethics. Psychology is both the peak of the natural sciences and the basis of ethics. As the whole conception of the world (*Weltauffaussung*) is mirrored authentically in each single part of the system, so also each component of the macrocosmos, which consists of nature and the spiritual world, is completely mirrored in the microcosmos of the soul.

Further arguments for the central position of psychology come from more clearly epistemological considerations, which relate to the philosophical debates of Tredelenburg's time: as metaphysics was no longer able to shed light upon questions concerning knowledge, philosophers turned to psychology, the discipline tightly connected to the firm ground of experience. The task of metaphysics was still to explain and to found objectively the principles of psychology.

When such objective principles of psychology are considered in an internal perspective, they have a purely subjective value and explain merely what is primary to us, that is, to our subjective knowledge. But when we want to find an objective foundation of these principles, it becomes necessary to ask what is primary according to the nature of things. The universal principles of nature, which must be investigated by metaphysics, are prior to the principles disclosed by psychology, which explain the concept of soul and give the basis to the so-called sciences of the mind. In the study of the psychical processes and in general in the description of knowledge and ethics, it is necessary to take into consideration the emergence and genesis of the principles of the mind. Such considerations explain the central position of psychology within the genetic system of philosophy. It is obvious that logic, understood as the basic theory of knowledge and philosophy, must provide the strength and foundation of genetic knowledge. Nevertheless, only psychology is able to investigate knowledge in its process of becoming, in its genesis, and in this task psychology joins metaphysics, the investigation which alone can legitimately claim to reach to the heart of nature (which is originally creative) and thus to reconstruct reality in its objective structure.

From where does psychology get its notions and concepts? The concept, when considered formally, consists of two aspects: its object and its fundament. The object is the end result of intuitive knowledge which refers to what is sensibly present and extended (*sinnliche Gegenwart*, *sinnliche Breite*), that is, to a correlative term

which is determined both temporally and spatially. This holds for the objects of external sense (or in Brentano's terms, the objects of external perception). However, the specific ambit of psychology concerns other kinds of objects, the objects of internal sense, such as thoughts, pleasures, sorrows, temperament etc., whose main characteristic is that they lack extension. Unlike external objects, internal objects are not known through the combined contribution of different senses or different ways of knowing. The only proper object of psychological knowledge is action which flows in a temporal stream (we will later see that Brentano too saw temporality as the main characteristic of mental states and of the soul). The temporality of the psychic object constitutes a principle difficulty for the psychological investigation: its object is in a perpetual process of change.

In order to deal with this problem, psychologists have usually resorted to observation, and turned from particular phenomena to investigate the general character of the living ensouled being in its wholeness. In other words, only observation seems to be able to secure the necessary link to experiential data, and this makes it possible to avoid the difficulties of a purely metaphysical approach. But self-observation, or introspection, causes almost insuperable difficulties, since in the case of mental states, the observer and observed phenomenon coincide. One inscrutable and mysterious entity should be able to understand another object which is equally inscrutable and mysterious. In addition to this difficulty another emerges: observed phenomena are constantly fleeting and are affected by the individuality of each observer. For these reasons, experimentation is impossible; the observer and observed are not adequately separated or independent. Moreover, self-observation seems either to lack all distance necessary for observational activity or else to distort the original experience by introducing an irremediable split into psychic life between the observing activity and the observed objectivity. In the latter alternative, the original mental connection is replaced by a new one which is totally different, and thus the natural order of the observation fails. If we try, for example, to observe a passion in its natural course, we temper the passion to the point of annihilating it as what it is.

From an epistemological point of view, however, it is possible to use observation in a fruitful way, on the condition that the observed phenomena are taken as mere signs of underlying unknown causes. In other words, the investigation cannot leave the level of the *causae cognoscendi* and cannot pretend to penetrate into the real nature of things. This is generally true of all science, but in the case of psychology, a specific dilemma emerges, because we can make two different kinds of observations: observations of others and observations of ourselves. At first sight, observing others seems to be unproblematic, but this is complicated when we notice that we can understand and explain the mental states of others only by presupposing that we can experience similar states in the first person. We can represent states that are formed as permanent dispositions in others, even creatures very different from us, only because such states exist potentially in ourselves and have taken place, at least once, in us. This means that the observation of other minds presupposes self-observation, as past and as potential. As pointed out above, self-observation rarely captures the psychic phenomenon in its original features. Moreover, self-observation can be carried out objectively only by specially educated individuals.

The product of self-observation is a modification of the original state and as such is a more complicated state. Additional factors, such as habituation and repeated practice, or the interest and the motivation for making observations, contribute to the constitution of the phenomenon.

Measurement plays an essential role both in the natural sciences and in scientific psychology: we can give quantities of intensity to the activities of the soul and these activities can disclose remarkable variations that can be measured. Even though the activity as such does not have any extension, we need to resort to the order of extension if we want to use instruments of measurement. In psychology, measuring is bound to the circumstance and the temporality of the phenomenon; it is not possible to refer to any uniform object outside the measured phenomenon. A thought takes place at a certain temporal moment, and its measuring must be included at the very same moment.

Is it possible to resort to other means in order to determine the object of psychology? According to Trendelenburg, it is necessary to consider the soul in its twofold relation to nature and to the spiritual world. The soul does not disclose itself immediately and directly but only indirectly, through this double relation. We can assume with good grounds that Trendelenburg wishes to abandon all substantialist conceptions of the soul; most importantly, this is indicated by his rejection of rational psychology and the inheritance of Leibniz and Wolff. Thus, the soul must be conceived in functional-operational terms by reference to the categories of movement and purpose and their interrelations. To put it in other words, the soul is defined by its dynamic relations to both natural being and spiritual being, and it finds its own accomplishments in them.[2] Here Trendelendenburg's solution seems to anticipate Brentano's psychology of act. In Brentano's conception, the principal mark of psychic activity is intentional reference, which has the twofold character of dynamism and directedness toward an aim, that is, toward an immanent objectivity understood as the end point of psychic relation. In contrast, the activities of organic nature are aimless when considered apart from the animated whole to which they belong. They live in the soul and refer to the soul in the same way as means and methods refer to the aim that has to be realized. The whole spiritual world is mirrored in the soul: the manifold essence of the world is expressed in the plurality of the different operations of objectification which the soul accomplishes and which transcend the sphere of the individual and realize themselves on a communal and collective level.

In 1833 Trendelenburg published a commentary on Aristotle's *On the Soul* (*De anima*).[3] This interpretative work meant to Trendelenburg an important opportunity to confront Aristotelian psychology and to study its viability. The confrontation with Aristotle made a deep and lasting influence on his thinking and is clearly visible in his masterpiece *Logical Investigations* and its doctrinal elaborations. Following

---

[2] For this aspect cf. Gershon George Rosenstock, *F.A. Trendelenburg: Forerunner to John Dewey* (Carbondale: Southern Illinois University Press, 1964), 88–90.

[3] Friedrich Adolf Trendelenburg, *Aristoteles de anima libri tres: Ad interpretum graecorum auctoritatem et codicum fidem recognivit, commentariis illustravit F. A. Trendelenburg* (Jena: Walz, 1833; Berlin: Weber, 1877; Graz: Akademische Druck- und Verlagsanstalt, 1957 and 1964).

Aristotle, Trendelenburg studied the structure and function of the soul in its organic context, and this made it possible for him to take into consideration the soul's interaction with biological and environmental factors. He renewed Aristotle's understanding of sense-perception and the doctrine of the soul as the end of the organism in the light of two contemporary developments: nineteenth century psychology and physiology of sense-organs[4] and the organic conception of the soul which had its origins in Romantic philosophy of nature.

Trendelenburg's theoretical debt to Aristotle is evident in his definition of the soul as *entelecheia*, i.e. as a primary act of a natural organic body. The soul is here understood both as the origin of the body and as its end result. On the one hand, Trendelenburg, directly influenced and inspired by Aristotle, claims that the soul relates to the body in the same way as vision relates to the eye. On the other hand, he argues that the soul is the result of the dynamic development of the natural organism, understood in a genetic-evolutionary sense. Thus, the soul does not transcend the body, but is independent of the body insofar as it is the body's essence or nature and its product. Trendelenburg describes the functions of the soul by drawing directly on Aristotle's distinction between the vegetative, the sensitive and the intellectual soul. In the case of the vegetative soul, two functions can be distinguished: assimilation and reproduction, the former being the transformation of inorganic matter into organic matter. Animal life is characterized by the sensations and the functions of desire; but in human beings, volitional and cognitive functions, will and thinking, arise and these mark the transition from the organic level to the ethical one. The primary and immediate aim of every organism is self-preservation, and it is the function of desire to ensure this: the gratification of desire concurs with the organism's self-realization. Self-consciousness, which is independent of physiological processes, follows the sensation of self and appears together with the emergence of thought.

Trendelenburg draws on these topics in his lectures on psychology which he gave in Berlin between 1840 and 1870. His interpretation of Aristotle's *On the Soul* as well as his investigations into logic, epistemology and physiology merge with his elaboration of psychological theories.[5] He argues that the soul should not be understood as an isolated object but must be examined in its organic context, and in this context he emphasizes the two main attributes of the soul: its movement (the dynamic principle) and its directedness or relatedness to an end. Understood as a thought of a self-realizing end, the soul develops itself through different levels of being alive, i.e. the levels of vegetable life and animal life, that constitute the side of nature, and the level of human life that represents mind or spirit. The human soul is expressed in its basic activities of sensation, knowledge, desire and will. These are

---

[4] See, for example, Karl Ernst von Baer, *Über Entwicklungsgeschichte der Thiere* (Königsberg: Bornträger, 1828); Johannes Müller, *Handbuch der Physiologie des Menschen* (Koblenz: Hölscher, 1840); Gustav Theodor Fechner, *Elemente der Psychophysik* (Leipzig: Breitkopf & Härtel, 1860); Hermann von Helmholtz, *Handbuch der physiologischen Optik* (Leipzig: Voss, 1867).

[5] Trendelenburg was a close friend and fellow of J. Müller, the initiator of modern physiology in Germany.

said to correspond to the manifestations of the Ego (the Self), which is the whole that constitutes the ground for the different functions. Among these, desire is given the primary role as the fundamental activity, which assures the reciprocal co-ordination of the other activities and constitutes their dynamic principle.

This revival of Aristotelian psychology happens in the context of a heated debate on the epistemological consistency of the new science of psychology. Trendelenburg develops and defends a third way mediating between the opposites of rationalism and empiricism.[6] The rationalistic approach disregards the results of physiology and the sciences of nature in general, insofar as it only takes into account the essence of soul; the empiricist approach remains satisfied with a theoretical framework which tries to explain psychological events merely by resorting to observation and to mechanistic models, without providing any metaphysical foundation. The formal definition of the soul as the thought of a self-realizing end allows Trendelenburg to explain a wide range of psychological phenomena and activities and to account for their interaction within the environment by a unitary principle. Based on this definition, Trendelenburg describes a highly articulated system of psychic tools and ends which corresponds with his own ontology. The activities of soul are organized hierarchically in the very same way as reality itself is structured into a multiplicity of levels (physical, organic, ethical and psychological), each of which has to be investigated by a particular science. Metaphysics does not exclude empirical research; rather it makes empirical science possible, insofar as it provides the conditions for scientific cognition by investigating the formal structures and principles of being.

According to Trendelenburg, the other main character of the soul is movement. Movement is understood as the principle of reality in Aristotle's sense of efficient cause and final cause. In his *Logical Investigations*, Trendelenburg develops a Hegel-inspired dynamic conception of reality in order to explain its multiple articulations and the connection between its different constitutive levels. For this purpose, he needs to postulate an external movement of nature, taken in an ontological sense, as well as an internal movement of the mind, in Trendelenburg's own words, "the constructive movement of spirit." These two movements refer to each other in the same way as the matter of an entity refers to its form. Trendelenburg needs these two co-referential movements of nature and spirit in order to explain the activity of the soul and to secure the correspondence between thought and being; his explanation rests on their structural homology and on the idea that they belong to the same ontological genus.

The historical background of Trendelenburg's approach is again in Aristotelian physics and psychology: the experiential evidence of movement is the axiomatic starting point for both sciences. Trendelenburg reinterprets and reformulates this Aristotelian principle in his theory of knowledge by emphasizing the connections between perception, thought and imagination. Constructive movement is both the

---

[6] On the development of philosophical psychology in Germany during the nineteenth century, see the exhaustive reconstruction of Klaus Sachs-Hombach, *Philosophische Psychologie im 19. Jahrhundert: Ihre Entstehung und Problemgeschichte* (Freiburg-München: Alber, 1993); on Trendelenburg, see 183–93.

metaphysical principle of the mind and its original act: in the latter role, it must be understood as a creative disposition that characterizes knowledge in its development through all its stages. Trendelenburg starts from a critical examination of the way in which Kant tried to connect intuition and intellect – the levels of transcendental aesthetics and transcendental logic – by resorting to "the artifice of schematism," as Trendelenburg calls Kant's solution. He argues that Kant first postulated the separateness of these two sources of knowledge, regarding them as static entities, and then tried to restore their unity with inappropriate tools. Contrary to this Kantian approach, Trendelenburg proposes that we should start again from the constitutive continuity that characterizes knowledge in its genetic development and should focus on the common characteristics of its different components. Rather than inquiring into the validity-conditions of knowledge, he wants to reconstruct the process of knowing in its concreteness.

The first stage is sense perception which, according to the Aristotelian understanding, is defined as a special kind of movement or change, i.e. as the assimilation of the pure form of the object without its matter. Perception only appears to be a passive alteration, for the creativity of the constructive movement of the spirit permeates knowledge in its entirety, beginning from its first steps. The evidence for this is the relative independence that sensory activity has from corresponding impressions and the role played by imagination in the process of knowledge. Trendelenburg goes so far as to identify the constructive movement with imagination, in order to secure its creativity and to explicate its synthetic function. He thus combines Aristotelian psychology with Kant's transcendental deduction (as it is carried out in the first edition of his *Critique of Pure Reason*[7]), and deduces both the forms of intuition and the concepts of intellect from the a priori constructive movement of the spirit. In this reinterpretation of Aristotelian psychology he makes use of the achievements of physiology at that time.

Trendelenburg illustrates the connection between movement and imagination by pointing out that movement as such is not accessible to perception and that we can only know movement by its production. He contests the traditional conception that assigns movement to productive imagination. Aristotle had argued that imagination has the subordinate status of a passive and receptive faculty, depending on and deriving from intuition. Trendelenburg abandons this notion, and strives to strengthen the status of imagination. He conceives imagination as an autonomous activity of the soul and argues that it coincides with intuition and functions as the fundament of sense perception (in the Kantian sense of a possibility condition). Understood as productive fantasy, imagination gives a definite and complete form to the images of external objects, since it presides over the unifying synthesis of the elements, the impressions and the matter that the senses provide for the constructive mind. Here again Trendelenburg seems to draw on the Aristotelian distinction between form and matter and on the classical thesis that matter as such is

---

[7] Immanuel Kant, *Critique of Pure Reason*, ed. and trans. Paul Gruyer and Allen W. Wood, in *Works of Immanuel Kant*, vol. 8 (Cambridge: Cambridge University Press, 1998).

unknowable. On the one hand, he argues that movement produces the form of things and in this function constitutes the a priori of knowledge; on the other, he claims that matter is the necessary substrate of the representation of movement and experience and that, in contrast to form, it has to be determined.[8]

Trendelenburg's explicit reference to Aristotle's concept of *phantasia* allows a comparison between his conception of imagination as movement and the treatment that Aristotle gives to *phantasia* in *On the Soul*. Trendelenburg's commentary on this work proves crucial in the explication of his own view.[9] In the introduction to *On the Soul*, Trendelenburg emphasizes a difficulty that strains the reconstruction of Aristotle's concept of imagination: the topic is not discussed at length in any text. In Aristotle's *On the Soul*, there are only brief notes on this topic, and his other works do not include much discussion on imagination and its relation to the other components and functions of psychic activity. In *Rhetoric*, imagination is connected with pleasure. In *On Dreams*, Aristotle discusses the question whether imagination belongs to the realm of opinion and noetic thought or rather to the level of sense experience. It is well known that in the Aristotelian framework imagination is a "stateless" faculty, which oscillates between the aesthetic and the noetic, and does not have a clear status or an autonomous character. The only thing that is certain is that the faculties of the soul arrange themselves in a progressive sequence, from sensation to imagination and from imagination to intellect, but the nature of their mutual interaction remains problematic. Open questions include, for example, the conjunction of intellect with the other faculties and the relationship between active and passive intellect. On the one hand, Aristotle affirms that sensation and imagination are influenced by the intellect, and on the other hand, he insists on the separatedness of the intellect and attributes to it the task of knowing suprasensible being.

Trendelenburg tries to solve these difficulties by resorting to Kant's concept of transcendental imagination and by introducing his own theory of constructive movement into Kant's theoretical framework. We have already seen that Trendelenburg's understanding of movement owes a debt to the connection Aristotle made between the soul and movement. He argues that a majority of psychic states results from the alteration (*alloiōsis*) of the sensible parts of the soul: pleasure and pain derive from acting, memory and hope, which derive directly from the imaginative faculty. This explanation rests on the concept of movement, because alteration is one of the modalities of movement. In an analogous way, Trendelenburg connects imagination with memory and expectation by making a distinction between reproductive or figurative movement (*abbildende Bewegung*) and prefigurative movement (*vorbildende Bewegung*): the former takes up past and present sensible materials, and the latter projects in advance the sense data of present perceptions, according to a definite form. Both modalities of movement share a retentive function, since both integrate

---

[8] Friedrich Adolf Trendelenburg, *Logische Untersuchungen*, vol. 1 (Leipzig: Brockhaus, 1840; Hirzel, 1862 and 1870; Hildesheim: Olms, 1964), 252f.
[9] Trendelenburg, *Aristoteles de anima libri tres*, 141–43.

sense data into permanent forms, even if these forms are not given in the present; on the other hand, the very same forms impregnate the present perception.

In particular, the faculties of acting are referred to sensation, because these faculties are moved or affected by the actions exercised by the sensible being (according to Aristotle, imagination is a prerequisite for the capacity to act and to undergo actions). In addition to action, the faculties of memory and expectation are also referred to sensation. In contrast, the states of the noetic part of the soul are not alterations of any sort and are not generated.[10] The sensible object activates the sensory faculty which is only potential before this and as such neither suffers anything nor is altered, but is complete and perfect in itself. Alteration is a specific kind of movement; according to the canonical Aristotelian definition, it is an imperfect or incomplete act (*ateles energeia*). The soul pursues or avoids the object, depending on the pleasure and pain caused by the object, and this activity reminds the activity of affirming and denying, that is, the synthetic and diairetic functions of the judging intellect. When the soul feels pleasure, it acts in relation to the good through the sensible medium; and when it feels pain, it acts in relation to the evil.

Aristotle gives another argument that supports the interpretation according to which the connection between movement and imagination is close. When arguing against those who deny the existence of movement he refers to the incontrovertible proof of the senses. Even if it would be false to claim that there are some sort of beings in movement, even if this opinion would be a mere production of imagination, it would still be an indirect confirmation of the existence of movement, since both opinion and imagination are movements of the soul.[11] Aristotle certainly denies the identification of the soul and movement, at least in the sense of a full coincidence, but at the same time he affirms a connection between the soul and movement and also the dynamic character of psychic states.

To be sure, imagination differs from sensation and from thought, but it cannot exist without sensation, and its functions constitute an indispensable prerequisite of intellectual apprehension. In developed animals, imagination plays an analogous role to thought; it functions like a defective approximation. In contrast to this, genuine thinking is only possible for animals that have reason. Based on two considerations, Aristotle argues that imagination and intellectual apprehension should not be identified: (1) While imagination depends on the will of the agent, opinion is subject to the alternatives of truth and falsity. It must be noted however, that Aristotle's text allows a distinction between two different sorts of imagination: there is true imagination, which corresponds to the perceived entity, even in cases in which no perception is present or actual, and there is fanciful imagination, which lacks such a connection to perception and is free in its productions. In the latter case, the effective ontological consistency of productions remains problematic, since we cannot recognize their factual existence. (2) A present sensation affects the subject

---

[10] Aristotle, *Physics*, 247a1–b4, in *The Complete Works of Aristotle*, vol. 1, ed. Jonathan Barnes (Princeton: Princeton University Press, 1984).
[11] *Ibid.*, 254a24–30.

emotionally and directly, but a merely imagined sensation has no such power. This, however, conflicts with the fact that imagination is associated with hope and fear, and more generally with all psychic states which are connected to hope and fear.

The attempt to describe imagination proves extremely difficult, and it seems that the Aristotelian framework cannot help in the attempt to associate imagination with thought (which seems to include both imagination and intellectual apprehension) while trying to distinguish it from sensation. Imagination, when understood in the proper and non-metaphorical sense, produces in us an appearance (*phantasma*). This seems to indicate that it belongs among the faculties, by means of which we judge and hold something as true or false, such as sensation (even though Aristotle says in other texts that all sensations are true), opinion, science and intellect. Briefly put, according to Aristotle, imagination involves a discriminating function.

We can find a similar tension in Trendelenburg's account. On the one hand, Trendelenburg emphasizes resolutely that imagination cannot be reduced to thought and that it is similar to intuition; on the other, he recognizes in imagination a common underlying root that is shared by sensation and intellect, and conceptualizes this root as their condition of possibility. The latter notion is analogous to the tripartition presented by Kant in the first edition of his *Critique of Pure Reason*. Moreover, Trendelenburg identifies imagination and constructive movement: thought, as a reflexive and comparative activity, develops itself by an epigenesis from imagination.

10.2 Brentano and the Primacy of Psychology as the Fundamental Philosophical Discipline

Trendelenburg made an important contribution to the debate on the epistemological status of psychology and its object; Brentano reinterpreted and elaborated Trendelenburg's account in his own discussions on the same topic.

Brentano's scientific attitude was characterized by a strong aversion to all dogmatic and systematic philosophy, which derived from the severe criticism that his teacher launched against Hegel's dialectic. He too turned to Aristotelianism in his attempt to rehabilitate experience and individuality, and to defend them against any substantialist metaphysics and any abstract generalizations that sacrifice real beings to essences. Moreover, Brentano shared with Trendelenburg the conviction that the method of self-observation is inapplicable to psychology and that the purely experimental approach is inadequate in the investigation of psychic phenomena.

According to Brentano, there is a fundamental similarity between psychology and the other sciences of nature. The similarity comes from the fact that natural sciences investigate natural phenomena by resorting to methods which are characteristic of psychology, that is, perception and experience. Brentano's own early studies on Aristotle encouraged him to adopt an empirical approach in his investigations of the psyche. Here it must be emphasized that Brentano did not understand experience in the sense of the British empiricists, but rather in the sense of

pure experience stemming from Aristotle. Brentano's main aim was not to analyze any singular sense datum, but to reconstruct experience and psychic activity in its wholeness. Many contemporaries of Brentano, such as Lange and Fechner (and also Mach, in the wake of the latter), refrained from considering the soul as an essential entity and argued that the only task of psychology is to investigate single psychic phenomena and their mechanisms of aggregation, independently of their function in any psychic whole. In contrast to this, Brentano took a mediatory position: he argued that methodologically psychology consists of analyses of psychic phenomena, but that it nevertheless maintains a metaphysical and theoretical significance insofar as it postulates a unitary and synthetic activity of the psyche.

Brentano also shared another fundamental argument with Trendelenburg, who had presented a genetic reconstruction of knowledge in his *Logical Investigations*. Brentano inherited from his teacher an urgency to trace elementary concepts back to the intuition from which they originally derive, and tracing the conceptual syntheses back to the basic elements of which they consist. He argued that logic is founded on psychology in respect to these two forms of genesis. Trendelenburg had insisted that intuition has the key role of providing the basis for all conceptual thought in the genetic process of knowledge. In a very similar way, Brentano considered sensation as the origin of all psychic activity insofar as it marks the border between the psychic and the physical realm and has individual things as its objects, that is, real beings, which are specified by spatial and temporal determinations. As we saw above, Trendelenburg conceived sensory intuition as a mode of constructive movement, and claimed that it plays a mediatory role between thought and being, and between the psychic and the physical realm, and refers intentionally to *ens quam individuum*.

Brentano saw the task of psychology to be the investigation of the succession laws of psychic phenomena and the classification of their intrinsic elements. (These ideas comprise another link to Trendelenburg, who adopted an analogous approach in his manuscripts about psychology.) The first stage of the investigation belongs to genetic psychology, which uses inductive procedures; the second stage belongs to descriptive psychology, which works to classify psychic phenomena and, by following a deductive method, aims at an exact definition. However, Brentano's distinction between descriptive psychology (*Psychognosie*, in Brentano's own terminology) and genetic psychology is not clear, since psychic phenomena have a dynamic character and dynamic aspects which stem both from their categorical origin and from their functional role. This is true especially of the acts of representation (*Vorstellung*), which are understood as psychic phenomena that present something to consciousness (in this context, the term "phenomenon" has a neutral sense of the manifestation of a fact, not the Kantian sense which is opposed to "noumenon"). Representation is the principal psychic activity, insofar as it directs the formation of judgments, thoughts, and other movements of the soul. Psychic phenomena are characterized by evidence, but they also display intentionality, that is, an internal reference to an object (intentional in-existence of psychic phenomena).

Following Aristotle, Brentano too, like Trendelenburg, characterizes the psychic object as pure form without matter. Thus conceived, the object has "mental"

existence, but in this context, the word "object" does not refer to the thing as *something* perceived (and therefore does not imply any ontological reduplication of the thing), but rather refers to the thing *as* perceived. The mental object consists of two constitutive factors, the act of representation and of the represented object, and these are indissolubly bound together. Analogically to the distinction between the act and object, it is possible to distinguish, at the representational level, two constitutive parts of intentional reference: a primary object (for example, a physical phenomenon, a phenomenon of fantasy, and so on) and a secondary object (for example, the consciousness of imagining, the consciousness of seeing and consciousness of the other psychic states). It should be emphasized that Brentano's concept of intentional reference does not, and cannot, allow any difference between imagination and representation; a similar position was defended by Trendelenburg, who unified these concepts and defined imagination as representation (*Vorstellung*). Thus Trendelenburg's philosophy prefigures Brentano's theory of intentional reference: both conceived the object of representation as an individual being and conceived it in correlation with the activity of representing. This object presents itself as an image which is spatially and qualitatively determined and is the term to which the internal movement of the soul refers. The main character of psychic activity is its reference to something as an object and therefore its relatedness. This presupposes the primacy of the category of relation (*pros ti*, in Aristotle's terminology) as was established by Trendelenburg in the deduction of the categories presented in his *Logical Investigation*. Psychic reference, as distinct from relations in general, does not presuppose the real existence of its second relatum, the object; the only existence that is required is the existence of something that is able to think.

Brentano's attempt to reestablish psychology fits well in the context of the Aristotelian-Renaissance which was promoted in the first half of nineteenth century, mainly by Trendelenburg. The philosophical movements of the time were concerned with reflections on the methodological foundations of the sciences of nature and the sciences of the mind and they gave psychology the leading role as a fundamental philosophical discipline. The sphere of philosophy was understood to be more restricted than the sphere of the sciences, in the sense that philosophy was assigned the task of critical reflection on the basis of the sciences and on the role of a scientific theory of knowledge. This approach is directly inspired by Kant. Psychology took responsibility for the exigency of investigating the constitution of experience, scientific knowledge and even the validity of logic. From these discussions developed the diffusion of radically psychologistic tendencies, since psychology was here constituted as an autonomous scientific discipline.

According to Brentano, philosophy has to conform to the ideals of exactness characteristic of the natural sciences, though not in the trivial sense of merely transferring the procedures of natural sciences into its own field; and psychology has the task of providing philosophy with a methodological model – the call for evidence and for the exactness of inner perception and the descriptive approach. Brentano distinguishes natural sciences, metaphysics and psychology on the basis of their objects: the objects of the natural sciences are physical phenomena, the objects of psychology are psychic phenomena, and the objects of metaphysics are

suprasensible phenomena. Each of these sciences must have its own autonomous categorical and conceptual tools, in conformity with the character of their objects. In particular, psychology does not aim at formulating abstract theories of the essence of the soul or investigating its connection with the body; rather its task is to describe and classify the facts of psychic life according to their own nature. It proceeds on the basis of a strictly empirical approach, which does not require the abandonment of metaphysics, but on the contrary makes metaphysics possible as science.[12]

It is instructive, in this regard, to study Brentano's statements about the definition of philosophy and its delimitation from the other sciences. Following the main tendency of his times, Brentano distinguishes between science in a strict sense, which covers the ambit of the universal sciences, and science in a more general or broader sense, which covers the ambit of individual (or historical) sciences. The universal sciences are divided into two kinds: deductive sciences (mathematics) and empirical-inductive sciences. The latter include the sciences of nature (physics, chemistry, etc.), the sciences of the psyche and metaphysics. Their objects are respectively bodies, psychic phenomena and the universal determinations that are common to the physical sphere and to the psychic one. Philosophy is raised to the level of the universal sciences; it is understood as the universal science of experience, and this implies a crucial difference from Aristotle, who identified philosophy with *epistēme*, that is, with necessary and incontrovertible knowledge. Philosophy differs from the natural sciences, which concern external phenomena, because it investigates the ambit of inner experience and aims at establishing the universal laws that are valid in the psychic realm. The competence of philosophy extends, however, beyond the realm of the inner experience, insofar as philosophy also establishes laws that apply to those universal facts, which are common to the psychic and the physical spheres, i.e. metaphysical laws. Thus Brentano conceives of psychology and metaphysics as the two principal components of philosophy, and he reappraises the Ancient tradition which had taken Aristotle's *On the Soul* as a sort of second metaphysic.[13]

To the duality of the object of philosophy corresponds the duality of the concept of substance, which plays a key role in Brentano's reinterpretation of Aristotelian metaphysics. Objects of thought have a special ontological status: they do not exist apart from the relational activity of the knowing subject. This holds both at the level of sensible perception and at the level of the discursive intellect, which includes judgement operations. For Brentano, substance is the fundamental meaning of being to which all others determinations have to be traced back and which plays the leading role as the guiding principle in the "deduction" of the categories, that

---

[12] Franco Volpi, "War Franz Brentano ein Aristoteliker? Zu Brentanos und Aristoteles' Konzeption der Psychologie als Wissenschaft," in Kurt Feilchenfeldt and Luciano Zagari (eds.), *Die Brentano: Eine europäische Familie* (Tübingen: Niemeyer, 1992), 134–38 and 143–45.

[13] See Franz Brentano, *Unsterblichkeit* (manuscript LS 22/1, 1875/76) and *Alte und neue Logik* (ms. EL 108*/2, 1877) (Graz: Forschungsstelle und Dokumentationszentrum für österreichische Philosophie).

is, the highest genera of being. In the wake of Descartes' distinction between *res extensa* and *res cogitans*, Brentano distinguishes two kinds of substance: an extended bodily substance and a suprasensible substance which is merely spiritual (or cognitive, in current terms). Brentano insists on this duality right up to the latest phases of his thought, in which he integrates his ontological reflections into the framework of a continuum theory. The physical bodily substance is individuated by three-dimensional spatial continua, and the spiritual substance is individuated by the point on a non-extended temporal continuum without any dimensions. The temporal determination constitutes the ultimate difference that individualizes all things, both material and spiritual.

We notice, however, two important respects in which Brentano differs from Descartes. First and foremost, in Brentano's understanding, the duality of the bodily substance and the spiritual substance do not coincide with the duality of the inanimate and the animate: the soul is not simply a spiritual substance, but a whole which consists of two parts, a sensitive-bodily part and a cognitive-spiritual part. After *Psychology from an Empirical Standpoint* (*Psychologie vom empirischen Standpunkt*, 1874), Brentano argues explicitly that the soul is a composite but at the same time unitary substance, which includes psychic activities as accidents. This can explain how several different sensitive activities – even those which are very different in kind – can belong to one and the same consciousness: the evidence of perception is not possible without an essential unity with the perceived thing. The soul and its activities are connected not only in an ontological unity but also in a conceptual one, which is highly articulated. Brentano assumes that psychic activities are partially united and partially separate, but he also assumes that a similar partial unity holds between the activities and their subject. More precisely, psychic activities must have in common a substantial, cognitive-spiritual part, which guarantees their existence and which can subsist independently of them. The activities of the soul are therefore essential, but they do not exist autonomously, since they are accidents of a common substance, the soul, which gives them unity and makes them belong to one unitary consciousness.

The soul, when conceived in its individual particularity, is always and uniformly implied in every perception of its activities, since inner perception always refers to something individual, and not merely to a universal concept, though the soul can present itself only in association with something accidental, that is with psychic activities. Because the substance of the soul is individual, it cannot be identified with thought. Moreover, the determination that individualizes the soul is not accessible to thought. If the soul coincided with an act of thought, it would be something merely universal. We cannot isolate the universal concept of the soul from its individual differences; otherwise we would be able to look inside a foreign soul. On the contrary, it is true that we can perceive only our acts of thought. Even our psychic concepts are particular, according to the individual character of our own soul. Thus, analogy is the only way in which we can represent the psychic acts of other subjects, because we can grasp in their specific individuality only our own. There is, however, a common concept of the soul that we can share with others, namely the concept of the spiritual substance that we construct in analogy with the concept of bodily

substance. Bodily substances display specific determinations (quality, space), while the soul does not display any determination as such, either specific or individual.[14]

In his ontological investigations, Brentano tried – by an analysis of the categories – to trace the manifold meanings of being back to the principal meaning, namely to the fundamental category of substance. Brentano's interpretation of Aristotelian psychology belongs to this theoretical framework and thus it constitutes the main branch of his ontology: the object of psychology is the soul as substance, that is, as organized into a multiplicity of parts, in a similar way as being is organized into a multiplicity of meanings. The ontological model which directs the inner articulation of the table of categories also illustrates the relation that the soul has to the activities that specify it and refer to it, as well as the way in which accidents refer to the substance. Aristotle's identification of substance and unity is crucial to the affirmation of the fundamentally unitary character of the soul: if only *one* actuality and *one* form can belong to a matter or to a being, then this holds also, and especially, for the soul which is the substantial activity of the living being and gives unity to the parts of which the living being consists.[15] Generally speaking, it must be said that all parts are related by a separation, both material and logic: being divided into parts is a characteristic property of the living being, but not of the soul. The soul is not a substance as such, only in a derivative sense: only the living being, whose soul constitutes the actuality, is a substance in a proper and primary sense. If the soul cannot be defined apart from the living being, then it cannot be conceived as something divisible into parts: we can speak of the parts of the soul only in a metaphorical sense, insofar as the soul constitutes the living being and grants its unitary character as a organic whole. The soul is also the first principle of vital functions in their multiplicity.

Brentano draws on Aristotle's view that the parts of the soul are arranged dynamically according to a hierarchy of levels. Each level leads up to the higher level on the basis of their substantial affinity and, inversely, each level acts on the lower level, on the basis of the principle that the efficient cause of becoming (i.e. the passive substratum) potentially contains the final cause that directs the whole process and is actualised by the agent. The final cause presupposes the efficient cause, since it is actualized not by the agent as such, but rather by the agent insofar as he accomplishes a virtuality that is given in the passive substratum. To every part of the soul belongs a specific tendency and a specific way of acting; the soul parts do not constitute a heterogeneous multiplicity, but are mutually interrelated and constitute a dynamical whole. The principle of assimilation is already operative in each single part: the one who realizes the faculties of the soul is not the agent as such, but

---

[14] Brentano, *Von der Seele* (ms. LS 24*/2).

[15] Aristotle, *On the Soul*, 402b9 and 413b12, in *The Complete Works of Aristotle*, vol. 1, ed. Jonathan Barnes (Princeton: Princeton University Press, 1984); Franz Brentano, *Die Psychologie des Aristoteles, insbesondere seine Lehre vom* nous poiētikos, (Mainz: Kirchheim, 1867; Darmstadt: Wissenschaftliche Buchgesellschaft, 1960); *The Psychology of Aristotle*, trans. Rolf George (Berkeley: University of California Press, 1977), 42–6 (the page numbers refer to the original German edition).

the agent insofar as he acts according to the particular constitution of the soul as form. He actualizes what was in the faculties only potentially. In the vegetative soul, the form that actualizes the tendency of the passive substratum lies already in this subtratum and not in any further productions. The faculties of the vegetative soul are motor faculties and they produce something which is similar to the form that is immanent in them. The case of the sensitive soul and the intellectual soul is the inverse: the form is not in the subject of knowledge, but is in the object, as both the sensitive soul and the intellectual soul lack all form and all actuality.

This apparent poverty of the sensitive soul and the intellectual soul, in comparison with the vegetative one, manifests itself in an incomparable advantage: neither posseses in actuality any determinate form, and so both potentially possess all forms. For this reason Aristotle calls the sensitive soul and the intellectual soul "the form of the forms"; they are not unilaterally limited to any specific function.[16] Sensation requires the passivity of its subject in respect to the object: the subject and the object refer to one other in the same way as potential being and actual being. The process of sensation implies an initial dissimilarity between the subject and the object which disappears only at the end of the process when the subject assimilates with the object. This assimilation should be understood as a specific kind of alteration which is to be distinguished from alteration in the primary sense: the subject of sensation does not undergo any loss of form, but on the contrary, it actualizes a latent potentiality by taking up the form of the object. By drawing on medieval terminology, Brentano conceives sensation as a particular case of *alteratio perfectiva*, and this allows him to argue that the form of the perceived object is received objectively without any matter (and not only through the constitution of the sensory organ) and without any corruption process occurring in the subject of sensation.[17]

Another important topic that Brentano works out under direct inspiration from Aristotle's psychology is that of inner perception or inner consciousness, i.e. the kind of perception that does not refer to any external object, but to a psychic act. Brentano finds an antecedent of this idea in Aristotle's concept of 'perception of perception'. Aristotle discusses the question if this faculty belongs to a specific sense, specially assigned to it, or whether belongs to the proper senses or to the

---

[16] Aristotle, *On the Soul*, 432a1–4.

[17] *Ibid.*, 423b31–424a13. About the Aristotelian inspiration of the theory of the intentional reference, see Rolf George, "Brentano's relations to Aristotle," *Grazer Philosophische Studien* 5 (1978), 242–54 and "Einleitung," in Franz Brentano, *Über Aristoteles: Nachgelassene Aufsätze*, ed. Rolf George (Hamburg: Meiner, 1986), XIV–XVII; Volpi, "War Franz Brentano ein Aristoteliker?" 138–40; Dieter Münch, *Intention und Zeichen: Untersuchungen zu Franz Brentano und zu Edmund Husserls Frühwerk* (Frankfurt: Suhrkamp, 1993), 50. Other critics, like Richard Sorabji, affirm that Brentano's theory of intentional inexistence does not faithfully reflect the original intention of Aristotle: his physiological theory of the reception of the form without matter consists properly in underlining the assimilation of the sensory organ to the object and does not imply the moment of certitude and evidence that Brentano ascribes to the inner perception. See Sorabji, "From Aristotle to Brentano: The Development of the Concept of Intentionality," in Henry Blumenthal and Henry Robinson (eds.), *Oxford Studies in Ancient Philosophy. Supplementary Volume: Aristotle and the Later Tradition* (Oxford: Clarendon Press, 1991), 248.

common sense.[18] He is inclined to defend the first hypothesis: he argues that if the perception of perception did not belong to a specific sense, then it would be a property of an object of the proper sense or of the common sense. A sensory organ can only perceive determinations that belong to its sensory field, by virtue of its local unity, and cannot perceive the sensation about the difference between two different sensations which belong to different sensory fields. We can perceive such differences only through synthetic operations which cannot be executed by the specific sensory faculties that correspond to specific sensory fields. The perception of perception cannot be traced back to the local unity of the subject of sensation. Nothing prevents us, however, from resorting to the temporal unity of sensation in our attempt to solve the problem of the distinction between different sensory objects which belong to the same sensory field or to different sensory fields. The temporal instant – the presence – is the limit that both joins and separates the past and the future, in the same way as the point both joins and separates two lines. Brentano emphasizes the ambiguous status of (spatial and temporal) limits and argues that the limit is both unitary and twofold at the same time. The temporal instant separates two different representations and constitutes the transition from one sensation to another as well as the difference between the two. The temporal unity of the subject of sensation is not punctual or indivisible, like the unity of a mathematical point, but is extended in space.[19]

By developing Aristotle's conception further, and even forcing it beyond its limits, Brentano assigns a specific object to the activity of the sense of sensation. This object is the same present sensation, and Brentano characterizes it as "the internal movements of the sensitive part".[20] To be sure, Brentano identifies the discriminating sensory faculty (the "primary faculty", in the terminology of *On Memory*) with the Aristotelian common sense (Lat. *sensus communis*; Gr. *koinē aisthēsis*), and here he clearly follows Aristotle. But in addition to this, he must also differentiate the discriminating faculty from those sensory faculties whose objects are common sense objects, and, even more importantly, from proper sense objects. Besides these modalities of sense, Brentano discusses a third, which has the task of perceiving and discriminating not only sensations, but also the remaining activities of the sensory part, and which constitutes the first prefiguration of self-consciousness at the level of sensibility. This is Brentano's remarkable conjunction of the inner sense and the temporal determination, as suggested by the fact that he tracks the inner sense back to the temporal unity of the instant. In this way, Brentano superimposes onto the theoretical framework of Aristotle's common sense a concept that seems to be derived from Kant's transcendental aesthetics.

The subject of all senses, and the subject of the sense of sensation, is the living animated body, not the soul considered in isolation from the body. This subject, unlike the subject of intellect, can suffer alteration and corruption, because its form

---

[18] Aristotle, *On the Soul*, 425a14–426b26.

[19] Brentano, *The Psychology of Aristotle*, 85–92.

[20] *Ibid.*, 95.

is connected to matter. Thus the intellectual functions have a mental character and the sensorial functions have a bodily character. But the thinking and feeling subject does not lose its unity, despite this multiplicity of its psychic functions and faculties. Brentano concludes his analysis of the sensitive part of the soul with important remarks about imagination.

Aristotle defines the faculty of imagination as that which saves images from previous sensations without involving any present perception. Brentano proceeds by distinguishing between two kinds of images: images of those objects that belong to external perception and images of those objects that belong to inner perception (i.e. images of the acts of perception). He argues that the movements of sensible objects outlive the action that they exert during sensation and thus contribute to form a permanent disposition in the sense organs. It is this disposition that directs the mechanism of the association of ideas.[21] An association of ideas occurs when a present sensation recalls a past one, even when there is no direct temporal succession from one to the other and even when the two sensations belong to different contexts. The disposition that directs association is the sensible imagination, which is opposed to the intellectual one. Brentano identifies it with passive intellect.[22]

In this context, it must be noticed that Brentano insists on the structural parallelism between sense and intellect: both faculties have in common the necessary structural fact that something acts on them, which allows the transition from a merely potential disposition to real and actual knowledge. In sensibility, the role of the actor is played by external objects, not by the sensible qualities that potentially lie in the sense organs. The same structure can be found in the intellect: it too is essentially a potential faculty and needs the contribution of another active principle, i.e. the agent intellect, in order to reach actual knowledge. If the transition to actual knowledge could occur by means of the qualities of the intellect alone, then the intellect would not be able to know anything which differs from it. As pure actual knowledge it would be able to think only about itself, and would be some sort of pure self-consciousness. In Brentano's understanding, there is no contradiction between the fact that the intellect is immaterial and belongs exclusive to the soul (and not to the living body), and the fact that its constitution, like that of the senses, is principally passive. This means that the potential and passively constituted intellect is the only cognitive faculty that humans can have, while the active intellect is divine and supernatural.

How then can Brentano combine the immateriality and spirituality of the intellect with its necessary dependence on sensible representations, which constitute the precondition for the functions of the intellect? The claim is that every intellectual operation which is directed to the universal needs the support of a corresponding sensible image, like in the Aristotelian example of the surveyor who draws a triangle on the sand in order to calculate the sum of its angles. Even without an external

---

[21] Aristotle treats this topic in *On Memory*, 451b20, in *The Complete Works of Aristotle*, vol. 1, ed. Jonathan Barnes (Princeton: Princeton University Press, 1984).

[22] Brentano, *The Psychology of Aristotle*, 95–104.

material support, we always resort to images that have already been registered in our inner "tabula"; we need such images to abstract the intellectual concepts from them, concepts which subsist in them, at least potentially. This account explains how imagination is necessary for intellectual apprehension, and it demonstrates the structural homology between sense and intellect, both having the character of a potential faculty. Another justification for the homology between sense and intellect is the fact that the activity of judgement, whose operations are the *synthesis* and the *diairesis* (i.e. conjunction and separation), is already anticipated by the activity of the faculty of desire, which in Aristotle is strictly bound to imagination. At this level we notice an analogous binary structure: we desire or avoid the objects that we perceive and thus we produce connections and separations of sensible representation. The same occurs at the level of passive intellect: it does not receive the intellectual forms directly from the sensible objects, but rather abstracts them from the images that are first elaborated by the imagination; but unlike sensible perception, the intellect is not bound to what is immediately present in space and time.

So imagination produces a preliminary abstraction before the intellectual abstraction. Its operation consists in leaving out the spatial-temporal determination, in order to institute a virtual presence, a sort of deferred temporality, in the inner space of soul. The operation directs the cognitive processes of sensibility as well as those of the intellect, since the sensible apprehension also implies the moment of formal abstraction. This type of change or movement in general is called alteration, and it is an essential theme for Trendelenburg and Brentano as well as for Aristotle. It traverses the levels of knowledge and gives them all a dynamic character, since each cognitive act is an actualization of a potentiality which, like movement, is always coming to an end. The intellect does not know any image independently of sensible matter, but knows separately, through an abstraction, something that is intrinsically rooted in matter. The only exception to this process is self-consciousness and the universal concepts of the Ego (both understood in a completely different way as in Aristotle's account) which seem to belong to the suprasensible sphere and present themselves as something original without any connection to matter.

Brentano points out that one could raise a twofold objection against the argument that the intellect depends on images: if there were a dependence, then the intellect – the highest faculty of the soul – would be both subjected to the rule of imagination and conditioned by the actions of the bodies, and both implications compromise the immaterial character of the intellect and its independent existence in relation to the bodily sphere. Brentano tries to repudiate the first implication by referring to the intellectual motor faculty, and he tries to deal with the second implication by referring to the role of *nous poiētikos*.[23]

In respect to the intellectual motor faculty, which Aristotle postulates in analogy to the sensible motor faculty, Brentano remarks that this faculty differs from the will and from the intellect, even though it is bound to both and entertains with them the same relation of the bodily motor faculty to the other sensible faculties.

---

[23] Brentano, *The Psychology of Aristotle*, 145–54.

The accentuation of the primacy of movement as an active principle, which joins the sensory soul and the intellectual soul, is very similar to the one that we find in Trendelenburg's interpretation of Aristotelian psychology.[24] The dynamical approach of Trendelenburg's *Logical Investigations* also characterizes the psychological and metaphysical views of Brentano. The intellectual soul acts on the sensory functions and on the imagination, as is shown by the phenomenon of reminiscence (*anamnēsis*), which is analogous to deductive thought and belongs only to human beings, unlike memory (*mnēmē*).[25] Moreover, the intellect does not act directly on sensibility, but only through imagination, and the movements that it exercises continuously on the images, which are the substitutes of the sensible objects, are at the level of intellect.

The mediating role of the imagination, and its connection to movement, is an important theoretical insight both for Trendelenburg and Brentano. Imagination performs this task in a twofold way, namely both as sensible imagination and as intellectual imagination: it anticipates the higher intellectual activities already at the level of sensibility and, inversely, reacts on sensibility by virtue of the impulse that comes from the intellectual motor faculty. The sensible imagination sets sensible images (*aisthēmata*) in motion and directs them toward sensible desire, by the contribution of the sensible motor faculty. Intellectual imagination transforms the representations produced by the sensible imagination and anticipates the cognitive operations of the intellect, such as abstraction and judgement. Moreover, it has the task of disciplining the desires and the movements of the soul rationally.[26]

Thought does not occur without the contribution of images, which potentially contain intellectual concepts and which are the instrumental cause of thought. The active intellect does not directly stimulate the receptive one to think, but only by having previously stimulated the sensible soul to act on the receptive intellect. Brentano assumes that there is a bodily organ on which the active intellect operates. He refers – not so clearly – to "the central organ of the sensitive part which contains the intellectual part of the soul"[27] without giving any further explanation. In a preceding footnote,[28] however, he refers to the unity of the intellect in order to emphasize the impossibility of the idea that the intellect is located in the bodily organs that direct the proper senses. He argues that this is impossible since the organs are manifold and separated and therefore not compatible with the unity of the intellect. Brentano suggests here the hypothesis that it is the acts of the intellect on the bodily organ that are responsible for the unitary sense of sensation, which was expressly identified by Aristotle with the common sense (*koinē aisthēsis*). The connection between the common sense and imagination is demonstrated by the fact that the intellect needs images (*phantasmata*) in order to abstract the concepts. We have good reason to

---

[24] Trendelenburg, *Aristoteles de anima libri tres*.

[25] See Aristotle, *On Memory*, 451b20.

[26] Brentano, *The Psychology of Aristotle*, 159–63.

[27] *Ibid.*, 174.

[28] *Ibid.*, 152, n. 111.

suppose that Brentano speaks here of the primary sensitive faculty, in which the active intellect should be included as a disposition (*hexis*), namely as an accidental form and not as a spiritual autonomous substance, totally separated from the human soul. The active intellect is pure actuality and cannot suffer any alteration, neither from its own activity nor from the actions of the senses: what suffers alteration is properly only the passive and receptive intellect.[29]

The spiritual part of the soul is not included in the human body, but it belongs to the same reality, since sensation and thought must belong to the one and the same essence in order to be related. This reality is partly bodily and extended, and partly spiritual and without any extension. The spiritual part is dependent in its functions on the sensible part, because the apprehension of an universal concept presupposes that this concept is already given in a concrete and corresponding image. To argue that we grasp the concept in the images of fantasy is to say, that we are moved by the sensible-bodily part of the soul while thinking. Aristotle tried to solve the problem of the interaction between the bodily part of the soul and the spiritual part, by denying that this interaction occurs through a bodily force. If this were the case, then intellectual concepts too would have the same qualitative character as the sensory impressions. The influence of the bodily part on the spiritual part consists in the operation of the active intellect, which does not coincide with any thought, but plays the role of a precondition of every thought. Aristotle says explicitly that in this regard the spirit is an absolutely passive faculty, since it must be able to think everything: the famous comparison of knowledge with an empty wax table expresses this potential and virtual receptivity, since the soul in itself does not contain anything before it receives the fantasmatic images in which it recognizes the universal concepts through the action of the active intellect.[30]

Brentano draws on the results of his interpretation of Aristotle's psychology – especially the idea to the temporal unity of subject – in his later theory of temporality as the individuation principle of the mental states and of the soul.[31] He then argues that the soul constitutes an autonomous substance, even though it is strictly connected with bodily substance, since both constitute a continuum. According to Brentano, the temporal continuum must be considered higher, and it constitutes the most universal determination of beings. Whereas the soul is individuated only by the temporal determination, bodies are individuated both by temporal and spatial determinations and therefore constitute a minor ontological genus. This raises the crucial question of the relationship between mental and bodily substances: even if we suppose that the temporal determination is the common attribute of both ontological kinds, we still have to explain how the spatial determination relates to the temporal one. Brentano tries to answer this question by presupposing the existence of a four-dimensional, unitary continuum, whose fourth dimension is time, while

---

[29] *Ibid.*, 163–80.

[30] Brentano, *Das Buch Delta* (Ms. A 71/1). The famous metaphor of the wax table occurs in Aristotle, *On the Soul*, 430a1.

[31] On this topic, see Edoardo Fugali, *Die Zeit des Selbst und die Zeit des Seienden: Bewusstsein und innerer Sinn in Franz Brentano* (Würzburg: Königshausen & Neumann, 2004).

space constitutes its three-dimensional boundary. This solution is far from satisfactory, since it does not take into consideration the possibility that a pure temporal substance without any spatial determinations could exist (i.e. the soul) – or has to be conceived as existing. In addition to this difficulty, it can be asked if Brentano's theory can be purged from the dualistic remnants that afflict its original formulation and from the supremacy of evidence that it attributes to consciousness. To this end, it would be necessary to define in an unambiguous way an independent epistemic realm which would concern the investigation of consciousness and mental phenomena *iuxta propria principia*.

Chapter 11
The Problem of Mind and Other Minds in William James's Pragmatism

Sami Pihlström

Abstract The chapter explicates William James's pragmatist conception of the human mind and his way of approaching the problem of other minds. James's pragmatism is usually classified among empiricist and associationist philosophies of mind, but as is shown, it can also be understood according to its Kantian features. In James's view, the mind is really an active and a purpose-oriented organizing principle which structures our lifeworld. The main difference between James's pragmatism and Kant's transcendental philosophy is that James does not make any explicit distinction between psychological and philosophical inquiries into the mind; he based his philosophy of mind on the same introspective methods that he used in his psychological studies. The chapter shows that despite this difference there are interesting connections between the pragmatist and the transcendentalist traditions. The latter part of the chapter questions the recurrent accusation that James's philosophy is individualistic and even leads to solipsism. It is highly questionable, whether James's philosophy of mind really reduces the other person to something in my experience. James's position can be properly understood only if the problem of other minds is not merely understood as metaphysical or epistemological, but also, or even primarily, as an ethical question. Despite this, one can find certain egocentric traits in James's philosophy. These traits stem from the fact that James's notion of the ethical attitude is based on the requirement that the self must take responsibility for the other in order to correct its "instinctive blindness" to the other's goals and projects.

Keywords Pragmatism · self–other relation · solipsism · transcendental philosophy · ethics · James

American philosopher and psychologist William James (1842–1910) was, as is well known, one of the founders of modern scientific (empirical) psychology and also what one may call a "philosophical psychologist". The nature and functions of the human mind, consciousness, and the self were in the focus of James's intellectual

S. Pihlström
Department of Social Sciences and Philosophy, University of Jyväskylä

work in virtually everything he wrote, including his philosophical formulations of pragmatism and pluralism and his philosophico-psychological explorations of religious faith and experience. In particular, the importance of James's purpose-oriented philosophy of mind, which is sometimes labeled the "teleological theory of the mind", for his main philosophical position, pragmatism, is obvious. The core of pragmatism is the purposiveness of human thought. The *selective activity of thought* is, thus, the starting point of the pragmatic approach James defended (and provides also a central background for later pragmatist positions in metaphysics and epistemology). The mind is, James held, active in organizing the data provided by sensory experience; it is only in the context of such a selection, guided by our various interests and purposes rooted in our human practices, that our experience is *of* objects in the first place.

Thus, according to James, selective, purpose-oriented principles are constantly at work in our mental operations by means of which we structure the reality we live in. All descriptions of the world available to us – as the kind of practice-embedded creatures we are – are inevitably interest-relative. More metaphysically speaking, the objects we (selectively) experience *are* bits of experience, actively organized by the experiencing subject – which, however, is not a "thing" existing independently of the experience itself.[1] In brief, the mind is not a metaphysical substance or substrate of any kind for James; it is only with reference to their specific functions in the stream of experience that the elements of our mental activities can be distinguished from each other. Thus, we may say that, through his psychological and metaphysical writings, James established a "functionalist", dynamical and processual tradition in philosophical psychology and the philosophy of mind.[2] His philosophical psychology is *naturalist*, focusing on the natural functions of our minds, but *not* reductionist; James also had his spiritual and even mysticist moments.[3]

James spelled out the central ideas of selectivity, teleology, and interest-driven attention[4] in empirical detail in his psychological *magnum opus*, *The Principles of*

---

[1] As James once put it, "mind-stuff" is a kind of experience, while reality itself, "howsoever remote, is always defined as a terminus within the general possibilities of experience." Nothing, neither the mind nor the world "it" might be said to experience, transcends experience. (William James, *The Meaning of Truth: A Sequel to* Pragmatism [1909], in H. Burkhardt, Fredson Bowers, and Ignas K. Skrupskelis (eds.), *The Works of William James*, 19 vols. [Cambridge, MA and London: Harvard University Press, 1978], 75. My references to James's writings, unless otherwise indicated, will be to the *Works* edition [1975–88]).

[2] Of course, we cannot straightforwardly interpret James as a functionalist in the sense in which this word is used in contemporary philosophy of mind. But in a wider sense of the term, James was undoubtedly a functionalist, given his interest in discerning the functionality of various mental activities in their specific contexts.

[3] See, e.g., James, *The Varieties of Religious Experience* (1902 [1985]), and *A Pluralistic Universe* (1909 [1977]); I am not concerned with this aspect of his thought here. On James's mysticism, see, e.g., Richard M. Gale, *The Philosophy of William James: An Introduction* (Cambridge: Cambridge University Press, 2005).

[4] For a helpful discussion of James's notion of the "organizing principle" of "selective interest" and its significance for his methodology, see Charlene Haddock Seigfried, *William James's Radical Reconstruction of Philosophy* (Albany: SUNY Press, 1990), Ch. 3. A comprehensive treatment of

Psychology (1890), and adopted them as the psychological basis of his pragmatist views on meaning, truth, and experience in *Pragmatism* (1907)[5] and other later philosophical writings. The core of his view can be illuminated by referring to a well-known letter he wrote to Dickinson S. Miller in 1907:

> The world *per se* may be likened to a cast of beans on a table. By themselves they spell nothing. An onlooker may group them as he likes. He may simply count them all and map them. He may select groups and name these capriciously, or name them to suit certain extrinsic purposes of his. Whatever he does, so long as he *takes account* of them, his account, is neither false nor irrelevant. If neither, why not call it true? It *fits* the beans-*minus*-him, and *expresses* the *total* fact, of beans-*plus*-him.[6]

The world, that is, all by itself "spell[s] nothing"; only an actively selected organization of the world is a fact. Hence, as James put it already in the *Principles*, individual human beings inhabit worlds that are at least partially of their own making, as "each of us literally *chooses*, by his way of attending to things, what sort of a universe he shall appear to himself to inhabit."[7] There is no essential structure of the world as it is in itself independently of our contribution and purposes. In the *Principles*, James also argued as follows: "*There is no property ABSOLUTELY essential to any one thing. The same property which figures as the essence of a thing on one occasion becomes a very inessential feature upon another.*"[8] Purposes are involved,

---

James's psychological and metaphysical views on the mind, consciousness, and experience – from *The Principles of Psychology* (1890 [1981–83]) to his late writings – can be found in the first half of T.L.S. Sprigge's massive volume, *James and Bradley: American Truth and British Reality* (Chicago and La Salle, IL: Open Court, 1993). For some recent discussions of James's notion of consciousness, including his views on non-conceptual experience, see the special issue of *Streams of William James* 6:3(2004).

[5] William James, *Pragmatism: A New Name for Some Old Ways of Thinking*, in *Works* (1975); also published as a one-volume edition together with *The Meaning of Truth* (Cambridge, MA and London: Harvard University Press, 1978).

[6] Quoted in Seigfried, *William James's Radical Reconstruction of Philosophy*, 371; see also James, *The Varieties of Religious Experience*, 346, for the same example. Seigfried (*ibid.*) refers to this example as illustrating James's "reconstructed pragmatic realism," noting that "to assert that the world can be grasped only through the interests we bring to it does not deny that it exists independently of us. We do not create the world, but we do participate in creating the world as known. Objectivity is not given, but the result of certain subjective activities." Compare Seigfried's statement earlier in her book (90): "There are no 'things' [...] apart from my interest. This should not be confused with the idealist claim that my interests create objects *ex nihilo*." Cf. Hilary Putnam's discussion of the beans-on-a-table case in his "Sense, Nonsense, and the Senses: An Inquiry into the Powers of the Human Mind," *The Journal of Philosophy* 91 (1994), 445–517, reprinted in Hilary Putnam, *The Threefold Cord: Mind, Body, and World* (New York: Columbia University Press, 1999), 447–48; on the issue of ontological constructivism – of "creating the world(s)" – in Jamesian pragmatism, see also Sami Pihlström, *Pragmatism and Philosophical Anthropology: Understanding Our Human Life in a Human World* (New York: Peter Lang, 1998), Ch. 1.

[7] James, *The Principles of Psychology*, I, 401; see also Seigfried, *William James's Radical Reconstruction of Philosophy*, 85.

[8] James, *The Principles of Psychology*, III, 959–60; emphases in the original; also quoted in Russell B. Goodman, *Wittgenstein and William James* (Cambridge: Cambridge University Press, 2002), 148–49.

for instance, in our decisions to treat either physical or historical (or some other) properties of an object as relevant in the identification of the object. "*Classification and conception are purely teleological weapons of the mind*," James concluded, having noted that it is too seldom observed "how entirely the intellect is built up of practical interests."[9]

This position, developed in more detail in *Pragmatism* and in James's other later works (and in the works of his pragmatist and neopragmatist followers),[10] amounts, in my view, to a much more deeply Kantian than a traditionally empiricist (or associationist) picture of experience, even though James is usually classified as belonging to the empiricist tradition of philosophy and philosophical psychology instead of Kantianism.[11] The Kantian features of James's thought will occasionally come to the fore in the following discussion, especially in relation to the issue of the *individuality* of what I (unlike James) like to call the "transcendental mind", or, more specifically, the issue of solipsism. This chapter, however, does not offer a historical exploration of James's debt to Kant (or any other classical philosopher). Nor am I primarily concerned with James's psychological theories, even though his conception of the mind – of his own, mine, and the others' – is indeed my topic. But I shall address his views on what Kantians may call the transcendental conditions of experience, albeit in a heavily reinterpreted manner.[12] Particularly, I shall take up the problem of solipsism as applied to the existence of other loci of thought, or streams of experience – viz., the problem of other minds.

It will be suggested toward the end of the chapter that the chief achievement of James's pragmatic approach to the other minds problem is his *ethical* way of dealing with the issue, although detailed argumentation for this interpretation would require a more comprehensive discussion of James's applications of pragmatism in, e.g., the philosophy of religion.[13] My remarks on this topic will, however, hopefully provide some novel perspectives on the relations between philosophical psychology and such central philosophical disciplines as metaphysics and ethics.

---

[9] James, *The Principles of Psychology*, III, 961, 941; see also Goodman, *Wittgenstein and William James*, 149, 174–75.

[10] For Dewey's similar views, see, e.g., John Dewey, *The Quest for Certainty: A Study of the Relation between Knowledge and Action* (New York: G.P. Putnam's Sons, 1960 [1929]). See further the discussion of pragmatism as ontological constructivism in Pihlström, *Pragmatism and Philosophical Anthropology*, with comparisons between James and such more recent pragmatists as Putnam and Nelson Goodman. R.B. Goodman compares James's anti-essentialism and anti-foundationalism with Wittgenstein's – while resisting the thesis that (according to either of these thinkers) we just "make up the world" (*Wittgenstein and William James*, 206).

[11] On the Kantian features of James's pragmatism, see Pihlström, *Pragmatism and Philosophical Anthropology*, Chs. 5–6; for more general discussions of the possibility of interpreting pragmatism in a transcendental (or quasi-transcendental) manner, see Sami Pihlström, *Naturalizing the Transcendental: A Pragmatic View* (Amherst, NY: Prometheus/Humanity Books, 2003).

[12] As suggested in some detail in Pihlström, *Naturalizing the Transcendental*.

[13] See, for some related reflections on these issues, Sami Pihlström, "William James on Death, Mortality, and Immortality," *Transactions of the Charles S. Peirce Society* 38 (2002), 605–28; Pihlström, "On the Reality of Evil: A Jamesian Investigation," *Streams of William James* 4:2 (2002), 12–21.

11.1 The Individual Mind and Its World

We may proceed by taking a look at how James – in his philosophical, rather than psychological, writings (overlooking the fact that no such distinction is readily applicable to his *oeuvre*) – defines "mind". In the context of his radical empiricism, the theory of "pure experience" as the ultimate reality, James characterizes the "mind" or "personal consciousness" as "a series of experiences run together by certain definite transitions," finding the "objective reality" as "a series of similar experiences knit by different transitions."[14] Accordingly, experience is, metaphysically speaking, prior to both subjectivity and objectivity.[15] It is important for James that "there is no bedding" in radical empiricism: no metaphysical or epistemological way to appeal to, say, an absolute mind or a Cartesian-like mental substance: "it is as if the pieces [of experience] clung together by their edges, the transitions experienced between them forming their cement."[16] In contemporary epistemological jargon, James's position is, clearly, coherentist rather than foundationalist, although he offers no technical analysis of knowledge or justification in terms of coherence.

In the fifth lecture of *Pragmatism*, James lists the mind or, rather, minds (along with, e.g., bodies, things, one time, one space, causal influence, and "the real") as belonging to a network of commonsensical concepts by means of which we structure our experience and the world we find ourselves living in.[17] These "*fundamental ways of thinking about things are discoveries of exceedingly remote ancestors, which have been able to preserve themselves throughout the experience of all subsequent time*," he suggests.[18] There are many different kinds of conceptual systems, he admits, but no system has displaced common sense.[19] Common sense, James

---

[14] William James, *Essays in Radical Empiricism* (1912), ed. Ralph Barton Perry (Lincoln and London: University of Nebraska Press, 1996), 80. (This collection is also published in *Works* [1976]).

[15] For a somewhat unusual perspective on James's radical empiricism, see Bruce Wilshire, *The Primal Roots of American Philosophy: Pragmatism, Phenomenology, and Native American Thought* (University Park: The Pennsylvania State University Press, 2000). Wilshire connects James's view not only with phenomenology (this comparison has been made earlier) but also with Native American holistic (and even shamanic) views of embodied experience as "worldly", natural-spiritual, etc., thus arriving at a devastating critique of any dualism between the mind and the external world (see especially 74–79). James's theory of truth, Wilshire argues, should also be understood along these lines – thus (*contra* his critics) without any commitment to subjectivism (*ibid.*, Ch. 5). James is a "direct realist" (*ibid.*, 21), whose conception of experience as world-involving is prior to any subject/object dualism; personal or private experience presupposes a common world; accordingly, "my mind" and "your mind" are "abstractions that forget the community of minding in the actual world that binds us together" (*ibid.*, 69). Yet, according to Wilshire, James asks, "Who are we? What is the world?" (*ibid.*, 57) – so, is there some dualism left, after all, some difference between us and the world? I shall return to what might be labeled the ineliminability of subjectivity toward the end of this chapter.

[16] James, *Essays in Radical Empiricism*, 86.

[17] James, *Pragmatism*, 85.

[18] *Ibid.*, 83 (emphasis in the original).

[19] *Ibid.*, 84. As a committed pragmatic pluralist, James wrote: "Common sense is *better* for one sphere of life, science for another, philosophic criticism for a third; but whether either be *truer*

tells us, is "a perfectly definite stage in our understanding of things, a stage that satisfies in an extraordinarily successful way the purposes for which we think."[20] Accordingly, the category of minds is among those *denkmittel* (as he occasionally puts it) that satisfy the purposes of our thinking; it is one of the selectively organized conceptual interpretations of experience. Yet, at the same time, it is this very mind or these minds that are supposed to do the selective, attentive thinking and to be guided by the purposes at hand. This creates a self-referential problem which I shall take up in due course.

Meanwhile, in order to place James's views in the tradition he mainly belongs to, we should note that his pragmatism has often been seen as radically individualistic, especially when compared to Charles S. Peirce's (as well as John Dewey's or Josiah Royce's) more socially oriented pragmatism(s).[21] The partial conflict between Peirce and James focuses on the tensions between realism and nominalism, antipsychologism and psychologism, as well as communal and individual accounts of inquiry and of the pragmatic maxim.[22] James, though by no means a straightforward nominalist in any traditional sense, focused on individual subjects' particular interests and experiences and on particular consequences of actions in his interpretation of pragmatism, while Peirce emphasized more general – more regularized, patterned, recurring – habits of action. Peirce's "extreme scholastic realism" admits "real generals" (universals, dispositions, laws, habits) more explicitly – more metaphysically – than James's nominalistically biased pragmatism.[23] James's admission

---

absolutely, Heaven only knows." (*Ibid.*, 93.) The point is, of course, that we have little use for the notion of truth in an absolute sense; in a more modest pragmatic sense, all of these diverging conceptual systems may offer us important "truths" about the world. I shall not dwell on James's much-debated and often misunderstood pragmatist theory of truth here (cf. *ibid.*, Ch. 6, and James, *The Meaning of Truth*, passim; see also Pihlström, *Pragmatism and Philosophical Anthropology*, Ch. 6, and especially Harvey Cormier, *The Truth Is What Works: James, Peirce, and the Seeds of Death* [Lanham, MD: Rowman & Littlefield, 2001]). Goodman (*Wittgenstein and William James*, 24–5) compares James's commonsense beliefs with Wittgenstein's "world-picture", which provides the background for distinguishing between truth and falsehood and also for making sense of both knowledge and doubt. Goodman admits that there is "both a Kantian and a pragmatic flavor to" James's list of commonsense concepts (*ibid.*, 26); I agree with this way of putting the matter. (The relevant reference to Wittgenstein here is to his *On Certainty*, trans. Denis Paul and G.E.M. Anscombe [Oxford: Blackwell, 1969].) On James's "ontological relativism", the view that we are entitled, "as our interests and purposes change, successively to take different worlds to be the real or actual world without inconsistency," that is, that what is real "depends upon the purposes and interests that are freely selected by a self," see Gale, *The Philosophy of William James*, 5.

[20] James, *Pragmatism*, 89.

[21] Royce's criticisms of James will be dealt with in some detail below.

[22] Cf. Sami Pihlström, "Peirce's Place in the Pragmatist Tradition," in Cheryl Misak (ed.), *The Cambridge Companion to Peirce* (Cambridge: Cambridge University Press, 2004), 27–57.

[23] Cf. Charles S. Peirce, *The Collected Papers of Charles Sanders Peirce*, 8 vols., eds. Charles Hartshorne and Paul Weiss (vols. 1–6) and Arthur W. Burks (vols. 7–8) (Cambridge, MA: Harvard University Press, 1931–58), 5.470 (c. 1906). Yet, James (*Pragmatism*, Ch. 5) does include the category of kinds as one among the commonsensical concepts that have succeeded in structuring our thought for ages, and continue to do so. See Cormier, *The Truth Is What Works*, for a more nuanced picture of James than the one usually found in general presentations; according to Cormier, James

of general ideas was psychologistic in the sense that generals were, for him, human beings' interest-relative classifications of reality through their practices of coping with it, instead of being (as they presumably were for Peirce) something ready-made in reality itself, quite independently of pragmatic interests and purposes.

As a corollary of James's insistence that abstract ideas must be put to work amongst actual experiential facts, we get his interpretation of the "pragmatic maxim", the pragmatists' major methodological tool: we should look and see what (conceivable) *particular* results would obtain, if this or that philosophical (or non-philosophical) conception of the world were true (or were to be held as true). Peirce consistently emphasized that it is in the "*conceivably* practical bearings" that "the entire meaning and significance of any conception" lies,[24] holding that the practical effects of our beliefs need never be actualized. For James, it was essential that the results *will* actually be reached at some point, however remote, in the course of experience. This is, perhaps, only a difference in emphasis, but a difference not to be neglected. For James, individual actions and their specific outcomes were what ultimately mattered; for Peirce, socially established habits and conceivable though perhaps never actualized effects were more important. In Jamesian pragmatism, the mind that selectively organizes the experienceable world is, then, an individual mind, not (say) the scientific community as an abstract whole.

Some relatively clear oppositions between Peirce and James can, then, be seen as connected with their different formulations and applications of the core of pragmatism, the pragmatic maxim. Peirce's scholastic realism, emphasis on community, antipsychologistic view of logic and emphasis on pragmatism as a logical principle (as contrasted to a speculative philosophical world-view, or *Weltanschauung*) conflicted – at least according to the received view – with James's nominalism, individualism, psychological (rather than logical) orientation, and psychologistic interpretation of pragmatism. Among these two founding fathers of pragmatism, James was surely the philosophical psychologist *par excellence*; for Peirce we may reserve the title "philosophical logician". Accordingly, the philosophical problem of the mind-dependence of the world (as experienced or cognized) is more urgent in James's pragmatism, though Peirce cannot entirely avoid this issue, either.[25]

---

also accepted habituality and dispositionality in his pragmatism. Cf. Seigfried, *William James's Radical Reconstruction of Philosophy*, 267, on James's pragmatic postulation of the modality of possibility and of "general rules". In *Pragmatism*, James argued that we should tolerate "abstractions" inasmuch as we "get about among particulars with their aid and they actually carry [us] somewhere" (40; cf. also *ibid*., 64). For further discussion, see Pihlström, "Peirce's Place in the Pragmatist Tradition"; Peirce's argumentation in favor of scholastic realism is considered (with lots of references to scholarly literature) in Pihlström, *Naturalizing the Transcendental*, Ch. 3.

[24] See Charles S. Peirce, *The Essential Peirce*, 2 vols., ed. The Peirce Edition Project (Bloomington: Indiana University Press, 1992–98), vol. 2, 145 (1903).

[25] On this problem in the work of these (and other) pragmatists, see Pihlström, "Peirce's Place in the Pragmatist Tradition," as well as Sami Pihlström, *Structuring the World: The Issue of Realism and the Nature of Ontological Problems in Classical and Contemporary Pragmatism* (Acta Philosophica Fennica 59, Helsinki: The Philosophical Society of Finland, 1996).

These conflicts and differences in emphasis are not unrelated to the ways in which these philosophers viewed the highly central notion of practical bearings or consequences of beliefs. For instance, in a note added in 1893 to his famous 1878 paper, "How to Make Our Ideas Clear", Peirce remarked that the maxim, understood as an application of the Biblical rule, "Ye may know them by their fruits," ought to be interpreted collectively, not individualistically.[26] The emphasis on the collective nature of the scientific enterprise, and of the habitually evolving rationality that human action manifests, extends through virtually everything that Peirce wrote. The individualistic overtones of James's pragmatism were as alien to him as James's psychologism. These divergences seem to come out especially clear when one takes a look at Peirce's 1903 Lowell Lectures on pragmatism, one of the most significant manifestations of his way of resisting the psychologization of his logical maxim of pragmatism. He remarks that his own early formulations of the 1870s were *too psychological* and that he no longer considers it satisfactory "to reduce such fundamental things [as the pragmatic maxim] to facts of psychology," because "man could alter his nature."[27] In these lectures, Peirce was concerned with demonstrating the truth of pragmatism – a demonstration that was needed, because pragmatism was (as already indicated) a logical maxim, not a "speculative doctrine". For James, on the contrary, pragmatism was the *prior* philosophical starting point, something against the background of which (only) the true significance of (say) logical demonstrations was to be assessed; pragmatism, moreover, was crucially linked to the psychological groundwork established as the theory of selective interests and the teleological nature of thought.

In addition to this brief comparison with Peirce, the philosophy and psychology of religion provide an illuminating example of James's psychologistic individualism. The individual believer, the subject of religious experiences, is, for James, clearly primary in relation to the church, the religious community, or the tradition within which those experiences usually take place. As James argues in great detail in one of his major works, *The Varieties of Religious Experience* (1902), the individual believer's experiences are the real locus of religion. Theological conceptualizations are secondary, if relevant at all. Charles Taylor's recent critique of this position is worth mentioning here: James may, Taylor suggests, simply have been too individualistic in his affirmation of the primacy of personal experience and in his corresponding neglect of the communal (and theologically interpretive) features of genuine religious life.[28] Some religious experiences may not even be intelligible

---

[26] Peirce, *Collected Papers*, 5.402n2.

[27] *The Essential Peirce*, vol. 2, 140.

[28] Charles Taylor, *Varieties of Religion Today: William James Revisited* (Cambridge, MA and London: Harvard University Press, 2002). Gale, speaking about James's "overglorification of the isolated individual" (*The Philosophy of William James*, 193), seems to agree. For other recent reactions to the *Varieties*, also discussing the tension between the individual and the community in religion, see the essays collected in Wayne Proudfoot (ed.), *William James and a Science of Religions: Reexperiencing The Varieties of Religious Experience* (New York: Columbia University Press, 2004).

in purely individualist terms. Yet, one may point out that James's position is first and foremost an affirmation of *moral* individualism:[29] he draws attention to the individual differences people are able to make in the world's scheme of things and holds that ethical values are ultimately based on individual interests. It is, arguably, *this* individualism that leads to James's profound respect for the pursuits of *other* individuals, and hence ultimately to a community-oriented, relational view of the good life (at least potentially including religious life). Even so, James's emphatically psychological approach to religious phenomena may eclipse socio-historical and theological approaches.

I shall not dwell on examples drawn from the philosophy or psychology of religion in this chapter, however.[30] I want to return to the question of the *mind-dependence* of all ontological structurings of the world, which I consider a key phenomenon in James's pragmatism (*contra* various kinds of metaphysical realism often assumed without further argument in contemporary philosophy). As reality gets conceptualized and thereby ontologically structured through the mind's selective attention, through our interests and, more generally, teleologically oriented activities – that is, since all objects are objects given to us in a context defined by human practices – the very identity and even the existence of the objects there are is (to some extent at least) mind-dependent. Alternatively, we might speak about *practice-dependence* here; it is only within purposively organized, interest-driven human practices that the mind *is* active in the object-constituting sense. Depending on the mind or on our practices, reality is *relative*; James is an ontological relativist.[31]

---

[29] Cf. here especially Ralph Barton Perry, *The Thought and Character of William James: Briefer Version* (New York: Harper & Row, 1964; the original two-volume edition was published in 1935), Ch. 25.

[30] See, for further elaborations, my above-cited papers, "William James on Death, Mortality, and Immortality" and "On the Reality of Evil". See also Sami Pihlström, "Mortality, Individuality, and Pluralism: William James's Democratic Religion," *American Journal of Theology and Philosophy* 26 (2005), 96–120.

[31] This, as observed above, is Gale's term (*The Philosophy of William James*, 5, 135–40). Gale goes on to examine and criticize at considerable length the tension resulting from James's "pragmatic self's" ontological relativism and his "mystical self's" absolute, non-relativized reality claims grounded in mystical experience (in which the very self to which reality is relativized, according to pragmatist ontological relativism, is "surrendered"); his main thesis is that James failed to unify his two "selves", the "Promethean" pragmatist self and the absolute mystical self. (*Ibid.*, 11, 188–89, 199, 219–22, 229–36; for his earlier, more scholarly treatment of this tension, see Richard M. Gale, *The Divided Self of William James* [Cambridge: Cambridge University Press, 1999]). I hope my reinterpretation of James in the next section goes some way toward a reconciliation, though as a response to Gale my proposal remains implicit; I cannot here discuss in any detail Gale's challenging arguments, and in any case I must set aside James's mysticist views here. For a recent criticism of Gale's interpretation, not denying the tension but invoking the possibility of integration (and contesting Gale's reading of James as a mystical absolutist), see James O. Pawelski, "William James's Divided Self and the Process of Its Unification: A Reply to Richard Gale," *Transactions of the Charles S. Peirce Society* 39 (2003), 645–56.

Minds themselves, and their activities, simply do not belong to a mind-independent reality "out there," structured prior to our selective interests and vital needs. The applicability of the category of minds – as well as of the equally world-relativizing category of "practice" – in the experienceable world is itself dependent on our *need* to apply it. We cannot settle the metaphysical questions first and wait and see how our commonsense ontology of the mental realm will be affected; we must, all the way from the start, take seriously the selectivity and purposiveness inherent in the mental lives we lead and be prepared to adjust our ontological conclusions (especially our conclusions regarding the reality of minds) accordingly. In James's memorable phrase, "the trail of the human serpent is [. . .] over everything."[32] We are "the very personages of the world-drama," not mere "readers" of the "cosmic novel."[33] What takes place within the "cosmic novel" narrating about the place of the mind in nature is, then, dependent on how we construe, and indeed construct, that novel. As James puts it in *Pragmatism*: absolute metaphysical reality, conceived as entirely independent of the human mind and thought, is "absolutely dumb and evanescent, the merely ideal limit of our minds"; thus, "[w]hat we say about reality [. . .] depends on the perspective into which we throw it."[34] "What shall we call a *thing* anyhow?" he asks a few pages later, and answers that this "seems quite arbitrary, for we carve out everything [. . .] to suit our human purposes."[35]

Now, the interesting self-reflective question we already briefly noted arises with full force as soon as this mind-dependence of ontological categories is applied to the ontology of the mind itself. The general issue of realism can thus be applied in the more local problem area of mentality. Is the nature of the (individual) mind, or its (our) mental states and activities, itself humanly constructed or mind-dependent, i.e., dependent on the human practices within which we treat the mind as a humanly relevant phenomenon and decide to commit ourselves to its reality, ontologically speaking? In a way, this question must be answered in the affirmative, if we are to follow James's line of thought. It is on the basis of their pragmatic functionality in individual thinkers' experience that we should evaluate our beliefs about the mind, including our beliefs about the selective-cum-teleological, interest-serving activities of the mind. Indeed, as I pointed out above, the postulation of the category of mind(s) has, according to James, turned out to be an enormously useful commonsense conceptualization of the world, something that scientific conceptualizations can possibly compete with but can presumably never replace.[36] Both common sense

---

[32] James, *Pragmatism*, 37.

[33] James, *A Pluralistic Universe*, 27.

[34] James, *Pragmatism*, 119, 118.

[35] *Ibid*., 122. I shall in a moment propose that these constructivist or even idealist formulations can fruitfully be interpreted in a Kantian-like transcendental fashion, yielding a pragmatic form of *transcendental idealism*. I shall not here elaborate on this point in any scholarly detail, though. For more comprehensive discussions, cf., e.g., Pihlström, "William James on Death, Mortality, and Immortality," and Pihlström, *Naturalizing the Transcendental*.

[36] As James would have little patience for recently influential eliminative materialisms and other physicalisms, I shall not pay attention to such scientistic doctrines here. (Cf. Pihlström, *Structuring the World* and *Pragmatism and Philosophical Anthropology*, for some reasons why pragmatists should reject such views).

and science are, however, conceptualizations of the world that the human mind itself produces. At this point, a vicious circle threatens. How could the very ontology of the human mind be metaphysically dependent on the selective activities and interests of that same mind?

As David Cooper puts it (in the context of Bergson rather than James), it would be incoherent "to identify the 'we' who 'carve up' reality into particular objects with individual flesh-and-blood people: for these would then be both carvers and carved."[37] Two questions arise. First, do we need to postulate a kind of meta-mind, something close to what James labeled the "*self of all the other selves*," an "innermost centre within the circle"[38] – perhaps something that James himself located in the "feeling of bodily activities"?[39] Secondly, is it coherent to suppose that the construction or "carving up" of reality is based on the activity (or feeling) of such a self?

11.2 Jamesian Pragmatism as Transcendental Philosophy

Whether or not James himself held the view, it seems to me that his pragmatism might be rescued from the *circulus vitiosus* charge by interpreting the object-constituting activity of the mind in a Kantian-like transcendental fashion.[40] The mind upon whose activities the ontology of the mind – the metaphysical picture underlying one's philosophical psychology – is dependent cannot simply be the empirical (psychological, factual) mind; it must be something analogous to what is labeled the "transcendental mind" (and what may perhaps be called "transcendental ego", "transcendental subject", or "transcendental consciousness", with minor differences in meaning) in the Kantian tradition – James's obvious distaste for such terminology notwithstanding.[41] Given James's adherence to a kind of phenomenological

---

[37] David E. Cooper, *The Measure of Things: Humanism, Humility, and Mystery* (Oxford: Clarendon Press, 2002), 92. Cooper reads pragmatists (particularly James) as "humanists" who draw on such metaphors of "carving up", "shaping", or "sculpting" the human world (*ibid.*, 96–8). Thus, he views James as a representative of the humanist tradition in philosophy which rejects the thesis that there is a (discursable, articulable) world in itself, independently of human concerns and projects. Cooper's complex argument turns on the suggestion that both humanism (even in its most developed form, viz., as "existential humanism") and its opposite, labeled "absolutism", are hubristic views that cannot in the end be maintained – and that, therefore, we should embrace a view according to which reality is ultimately a mystery (see *ibid.*, Chs. 11–13). Perhaps James's religious mysticism might also be accommodated within such a position, but this is not a topic I want to discuss here.

[38] James, *The Principles of Psychology*, I, 285.

[39] *Ibid.*, I, 288; see Goodman, *Wittgenstein and William James*, 93–4.

[40] Cooper (in his *The Measure of Things*) is not, in my view, sufficiently responsive to this option.

[41] Cf., e.g., James, *Essays in Radical Empiricism*, 86. As Seigfried notes, James "creates the paradigm of the full self" instead of resorting to a transcendental ego (*William James's Radical Reconstruction of Philosophy*, 362). James's anti-transcendental rhetoric is perhaps most explicit in *Pragmatism* and *Essays in Radical Empiricism*.

methodology in his "concrete" understanding of the human condition,[42] and given the transcendental orientation of (at least Husserlian) phenomenology, this reinterpretation of James's philosophical psychology might not in the end be as strange as it may initially sound.

I am *not* claiming that James postulated anything like the Kantian (or Husserlian) "transcendental ego"; on the contrary, he quite explicitly renounced any such entity, which he found mysterious. As Russell B. Goodman lucidly explains, the difference between James's and Wittgenstein's philosophical psychology (in this respect) is precisely that James sought for the "self of selves" scientifically and introspectively,[43] in a context "detached from the ordinary contexts in which we employ the word 'self'," whereas Wittgenstein's inquiry was " 'critical' in the Kantian sense," focusing on "the kind of answers we are able to give, rather than trying to give them," and on the illusions that tend to arise when words are employed outside their relevant language-games.[44] I by no means wish to downplay this difference between an empirical (introspective) and critical (Kantian or Wittgensteinian) orientation to the notion of the self, or the mind. But the point of the reinterpretation I am proposing is that the transcendentality of the mind (or the ego, or the self, or consciousness) ought to be located *in* the functional activities James took seriously in his empirical philosophical psychology, in the very selectivity and purpose-orientedness of the mind he saw as the basis of his pragmatism. This is to reinterpret not only James's philosophical psychology but the notion of the transcendental ego as well. The latter should be liberated from the fixed Kantian system of categories and turned into a more processual and dynamic – less stable and definitely non-Cartesian – principle to be found within our worldly practices of organizing experience.[45] In any event, the mind actively constitutes the world of objects, either within a commonsensical conceptualization or a scientific one, or under some other system of concepts that serves our interests. This activity in *Gegenstandskonstitution* is clearly something

---

[42] See, again, Seigfried, *William James's Radical Reconstruction of Philosophy*, and Wilshire, *The Primal Roots of American Philosophy*. For a comparison of James's and Wittgenstein's way of drawing attention to the concrete details of human life (and, correspondingly, to the "priority of practice"), see Goodman's illuminating analysis in *Wittgenstein and William James*.

[43] Goodman, *Wittgenstein and William James*, 96. On the relevance of introspection in the philosophy of mind and psychology, see Gary Hatfield's contribution to this volume.

[44] Goodman (*ibid.*) mainly discusses the second part of Wittgenstein's *Philosophical Investigations* (trans. G.E.M. Anscombe, Oxford: Blackwell, 1953) here; I shall not dwell on the issue of the exact relation between James and Wittgenstein, though. Goodman's main purpose is to show that, despite their major differences, James's and Wittgenstein's philosophical concerns were partly parallel and that James in fact crucially shaped Wittgenstein's thought on a number of issues. "James, in short, helped Wittgenstein understand some ways in which 'the human body is the best picture of the human soul'," Goodman concludes (*Wittgenstein and William James*, 118; the reference is to Wittgenstein, *Philosophical Investigations*, II, 178).

[45] See further Pihlström, *Naturalizing the Transcendental*.

that the Jamesian mind shares with the Kantians' and the phenomenologists' transcendental self.[46]

It is in this reconceptualized sense, then, that we may take James's philosophical psychology, despite its occasionally quite explicit anti-Kantian polemics, to be a form of "transcendental psychology". The true meaning of the transcendental is to be found within the empirical. Admittedly, the label, "transcendental psychology", may sound confusing, and we might easily avoid it, speaking about (say) "transcendental philosophy of mind" or "transcendental analysis of consciousness" instead.[47] But as James himself draws no principled distinction between philosophy and psychology, the one who attempts a transcendental rearticulation of his pragmatism should perhaps bite the bullet and use the (Kantian) expression, "transcendental psychology", in the rather non-Kantian, pragmatic way I am using it, being aware of the fact that this terminology may lead to confusion.

Even though my transcendental articulation of the Jamesian position is above all a *re*interpretation and a *re*articulation, doubts about the possibility of comparing James's thought to Kant's, let alone Husserl's, may of course still remain. A critic may legitimately wonder how the Jamesian mind can be properly seen as transcendental, if his psychology is first and foremost empirical. How can the analogy to Kant or Husserl be formulated in a manner that does not completely abandon James's own empiricist position? In particular, how *could* pragmatism and a transcendental philosophical framework such as phenomenology be brought together, given Husserl's and other phenomenologists' firm rejection of naturalism and psychologism (despite the rather similar emphasis on "concreteness" in James)? The transcendental ego in Husserlian phenomenology is not an entity at all (and *a fortiori* not the kind of substantial, "mysterious" entity James rightly renounced) but, on the one hand, an "empty pole" of intentional, cognitive acts and, on the other hand, a temporal structure of the stream of consciousness.[48] It is crucial that for Husserl and for many of his followers, a transcendental inquiry into the conditions for the possibility of experience is *eidetic*; accordingly, the conditions such an inquiry identifies are *essential* structures of experience. Furthermore, the way in which such eidetic structures are given is apodictically certain, and therefore phenomenology as a properly transcendental inquiry seeks to offer a foundation for all sciences.[49]

---

[46] For a number of perspectives on the possibility of reinterpreting the idea of the transcendental, see Jeff Malpas, *From Kant to Davidson: Philosophy and the Idea of the Transcendental* (London and New York: Routledge, 2003). None of the papers in that collection discusses James, though.

[47] Dan Zahavi suggested (in discussion) that the term "psychology" could be reserved to empirical study of the mind and the philosophical transcendental approach should be distinguished from such a study. See Zahavi's contribution to this volume, "Philosophy, Psychology, Phenomenology," which illuminates the phenomenological transcendental analysis of consciousness; among his numerous insightful treatments of phenomenology, see also Zahavi, *Husserl's Phenomenology* (Stanford, CA: Stanford University Press, 2003).

[48] This is not the right place to dwell on the different conceptions of the ego in the phenomenological tradition (e.g., the "egological" and "egoless" versions of phenomenology).

[49] These formulations are due to Sara Heinämaa (see also her paper in this volume).

Now, James (like most pragmatists later) will have none of this. His pragmatism abandons all kinds of foundationalism, claims to apodictic certainty, as well as essences.[50] His conception of philosophical methodology, moreover, rejects the claim (typical of phenomenology) to begin philosophizing from a pure, presuppositionless state by giving up the "natural attitude". James might argue (though he does not use these terms) that we can never really leave our natural attitude behind us; there is no way for us to analyze the eidetic deep structures of experience. We can never reach a purely philosophical, non-natural state in which a phenomenological reduction could effectively operate. There is no way to "start from the beginning", to begin philosophical thinking by bracketing all factual commitments, merely focusing on what is immediately given to pure consciousness; on the contrary, the pragmatist insists that we are inevitably tied to all kinds of factual and natural (pragmatic) commitments.[51] In this sense, there is an obvious difference between James's philosophical psychology and Husserlian phenomenology. Nor can James's position be straightforwardly assimilated to Kant's. Clearly, he gave up the Kantian table of categories, as well as the distinction between phenomena and noumena – and, more generally, he did not sympathize with Kant's complex architectonics at all.

Yet, as I have suggested, human practices, to which reality – or any humanly relevant real entity – is relativized (cf. above), can be seen as playing a transcendental role in James's (and other pragmatists') philosophy. It is through practices that things are identified and, to use a more phenomenological term, "constituted". Only in a purpose-oriented context established by a given habitual practice can a thing emerge as a thing for us. Practices may be compared to the "lifeworld" that Husserlian phenomenologists speak about, or, perhaps more plausibly, to Heideggerian "being-in-the-world". Moreover, in terms of James's radical empiricism, we may say that things are not simply the objects we (subjectively) identify and deal with; "we" and "they" are elements of the very same experience, which extends everywhere.[52] The lifeworld or practice within which our experience of objects is so much as possible is, for pragmatists as well as phenomenologists, a condition for the possibility of science, but pragmatists may treat this condition as more flexible and less stable, less apodictically certain, and more dynamically evolving than phenomenologists are usually willing to do. It can be an a priori condition only in a relativized and historicized sense.[53]

---

[50] Compare here the passage from the *Principles* (III, 959–60) already quoted above: "*There is no property ABSOLUTELY essential to any one thing.* The same property which figures as the essence of a thing on one occasion becomes a very inessential feature upon another." This is a clear rejection of eidetic structures.

[51] For a pragmatist rejection of the idea that a philosophical stance could be presuppositionless, see also Sandra B. Rosenthal, *Speculative Pragmatism* (Amherst: The University of Massachusetts Press, 1986), 201.

[52] James's notion of pure experience (elaborated at length in *Essays in Radical Empiricism*) is perhaps his most obvious link to phenomenology; however, this is the part of his thought that I (as I will suggest later) think should be subordinated to pragmatism, rather than *vice versa*.

[53] For a pragmatically relativized conception of the a priori, see Rosenthal, *Speculative Pragmatism*, and Pihlström, *Naturalizing the Transcendental*, Ch. 1.

While, in addition to emphasizing the transcendental status of practices, we may use the notion of a transcendental ego, as I have proposed, it is important to avoid postulating such an "entity" somewhere beyond experience. Despite their above-described important differences, both pragmatism and phenomenology can be seen as philosophical accounts of experience and of the mind that avoid such a postulation, developing a method that eventually turns back to a self-reflective examination of the ego or consciousness which is "immersed naively in the world,"[54] that is, to the very same natural (empirical) ego that can and must describe *itself* and its manner of experiencing objects both empirically and transcendentally. What this means is that we are (though perhaps always imperfectly) acquainted with ourselves *qua* transcendental selves already in our pre-philosophical everyday life, in the kind of practices that pragmatists like James draw our attention to. It is experience itself that requires us to be conscious of ourselves as not merely empirical but also transcendental subjects of experience. In my view, both transcendental phenomenology and pragmatism are, to adapt David Carr's formulation, "forever poised on the line between the natural and the transcendental standpoints."[55] Neither of these two philosophies claims that a transcendental mind somehow makes up the world *ex nihilo*; neither postulates two realms of things or egos (empirical or mundane and transcendental or other-worldly); for both, empirical and transcendental ways of describing the self in its constitutive activities are precisely two ways of conceptualizing one and the same experience, or stream of consciousness. As Dan Zahavi explains, the empirical and the transcendental subject, in Husserlian phenomenology, are not two different subjects but "two different takes on one and the same subject"; the difference is between two ways of conceiving of the subject – "as an object in the world" and "as a subject for the world, i.e., as a meaning-bestowing and world-disclosing subject of intentionality."[56] Again, I must emphasize that I am not claiming that James himself should be read as a transcendental thinker in this sense; what I am claiming is, more modestly, that a pragmatist employing his conception of purpose-oriented, pragmatic "carving up" of the world (and rejecting, on these grounds, metaphysical realism postulating a world in itself with its own essential structure) may interpret these ideas transcendentally, thus *both* avoiding the naively idealist "the mind makes up the world" line of thought *and* saying something more than the mere commonplace (accepted by most empirical psychologists) that the mind actively organizes experiences of the world. Such a reinterpretation may, and

---

[54] See David Carr, *The Paradox of Subjectivity: The Self in the Transcendental Tradition* (Oxford: Oxford University Press, 1999), 84, 97.

[55] *Ibid.*, 84. See also Pihlström, *Naturalizing the Transcendental*, 26–8, 87–9. In addition to Carr's work, I have found J.N. Mohanty's interpretations of phenomenology and transcendental philosophy helpful here: see Mohanty, *The Possibility of Transcendental Philosophy* (Dordrecht: Martinus Nijhoff, 1985), and *Transcendental Phenomenology: An Analytic Account* (Oxford and Cambridge, MA: Blackwell, 1989).

[56] Zahavi, "Philosophy, Psychology, Phenomenology" (this volume), Section 2. (Zahavi cites here Edmund Husserl, *Phänomenologische Psychologie*, Husserliana IX, Den Haag: Martinus Nijhoff, 1962, 294).

should, maintain and even emphasize the possibility of different perspectives on one and the same subjectivity, roughly in the way phenomenology does.

The Jamesian pragmatist may also hold that James *should* have developed his philosophical psychology in a more transcendental direction. It cannot be shown that James himself succeeded in formulating a transcendentally acceptable response to the transcendental problem of how the "carver" and the "carved" are related to each other. In my view, we should look for a promising answer in the direction of a synthesis of pragmatism and transcendental inquiry – a synthesis obviously differing from phenomenology in significant ways, perhaps hopelessly remaining within "the natural attitude" (from phenomenologists' perspective), yet perhaps similar enough to justify the brief comparison I have made and further, more substantial comparisons along these lines.[57]

There is also a metaphilosophical issue to be raised as another way of concluding the discussion of this section. Should (1) a psychological theory about the teleological and selective activity of the mind be used as the empirical basis of pragmatism, or should (2) pragmatism itself, rather, be taken as the ultimate – transcendental – philosophical (though fallible and revisable) ground of any empirical theories, yielding the pragmatic (mind-dependent), functional truth about the teleological and selective empirical mind? I am not sure how James himself would have answered this question (except that he would undoubtedly have avoided formulating his answers in any transcendental vocabulary). It should be clear, however, that my reinterpretation and comparison between Jamesian pragmatism and transcendental phenomenology supports the second alternative.

11.3 The Problem of Other Minds

At this point another problem – a thorny problem not only for James but for most, if not all, transcendental thinkers from Kant to twentieth-century figures like Husserl and Wittgenstein – arises. Having drawn attention to the individualistic emphasis of James's philosophical psychology, we may ask whether the transcendental mind whose activities constitute the experienceable world, including the psychological activities of the factual psychological mind, is necessarily individualist or even *solipsistic*. It seems that the transcendental mind is inevitably *mine* – or is *me*. So, we should examine whether Jamesian pragmatism, or the kind of transcendental

---

[57] As Zahavi puts it (*ibid.*, Section 3), it would be a mistake to believe that transcendental philosophy is "all one thing"; Zahavi's chapter is helpful also in highlighting the differences between Kantian and Husserlian conceptions of transcendental philosophy. According to Zahavi, transcendental phenomenology employs an "enlargened notion of the transcendental," approaching issues such as embodiment and intersubjectivity through a transcendental analysis. While the enlargement of the notion of the transcendental (and of the transcendental ego) to be found in (Jamesian) pragmatism is certainly different from the one Zahavi finds in phenomenology, it also welcomes the topics of embodiment and intersubjectivity as key elements of a picture of human experience richer than the one provided by orthodox Kantian transcendental philosophy.

psychology that (according to my rearticulation) arises as the true outcome of his empirically informed philosophical psychology, *can* really avoid solipsism. Doesn't the Jamesian investigation of the mind and its world inevitably begin with the individual subject's experiences and her/his practical action seeking to transform those experiences? Isn't the other only a pragmatically useful postulation *I* need in *my* practices, something that exists only in my experience (purposively selected and organized, and so on)? It is, I think, no exaggeration to say that James was the greatest American thinker preoccupied with the issue of solipsism at the turn of the twentieth century. Certainly the problem was for him much more vital than for the other classical pragmatists. Peirce and Dewey were simply too socially oriented to have taken the topic of solipsism seriously.[58]

It is not easy to find comprehensive or detailed discussions of solipsism in James, although *The Principles of Psychology* of course includes lengthy treatments of the notion of the self, the "stream of thought," etc. Later, the problem of solipsism did re-emerge in his writings. In a short essay first published in 1905 and later included in the posthumous collection *Essays in Radical Empiricism*, James tries to refute the criticism, presented by Boyd H. Bode in the same year, that his metaphysics of pure experience is solipsistic and cannot account for the "self-transcendency" of experience.[59] James's counter-argument can be seen as an application of the pragmatic maxim, which, we may recall, encourages us to look for the (conceivable or actual) practical outcome of the philosophical conceptions we are considering:

> Does it not seem as if the quarrel about self-transcendency in knowledge might drop? Is it not a purely verbal dispute? Call it self-transcendency or call it pointing, whichever you like – it makes no difference so long as real transitions towards real goals are admitted as things given *in* experience, and among experience's most indefeasible parts. Radical empiricism, unable to close its eyes to the transitions caught *in actu*, accounts for the self-transcendency or the pointing (whichever you may call it) as a process that occurs within experience, as an empirically mediated thing of which a perfectly definite description can be given.[60]

---

[58] A bibliography of pragmatism (John Shook, *Pragmatism: An Annotated Bibliography 1898–1940* [Amsterdam and Atlanta: Rodopi, 1998]) lists, however, several contributions to the debate on the relation between pragmatism and solipsism. During the heyday of the pragmatist movement, several critics charged pragmatists, especially James but also Dewey and the British pragmatist F.C.S. Schiller, of solipsistic consequences. James's replies to such accusations can be found in his two late collections of essays, *The Meaning of Truth* and *Essays in Radical Empiricism*. Sprigge – one of the few commentators addressing this topic – rejects solipsistic readings of both Bradley's and James's "panexperientialism" (see his *James and Bradley*).

[59] See James, "Is Radical Empiricism Solipsistic?" (1905), in *Essays in Radical Empiricism*, Ch. 9. James responds to Boyd H. Bode, " 'Pure Experience' and the External World," *The Journal of Philosophy* 2 (1905), 128–33. See also John E. Russell, "Solipsism: The Logical Issue of Radical Empiricism," *The Philosophical Review* 15 (1906), 606–13. A similar charge was raised by F.H. Bradley: see his letter to James (April 28, 1905), printed as an appendix in Sprigge, *James and Bradley*, 590–91. It is also possible to argue that the verificationism implicit in James's pragmatic conception of truth is not far enough from solipsism (cf. *ibid.*, 38).

[60] James, *Essays in Radical Empiricism*, 239.

For James, in brief, the transcending of the individual self in experience does not lead us anywhere beyond experience, because experience is, literally, everywhere. The world, according to radical empiricism, is (transcendentally) constituted by, or *is*, pure experience.[61] This *panexperientialism*, as Sprigge calls it,[62] is not to be equated with solipsism pure and simple, since experiences are not had or "owned" by any soul-like substance, nor constituted by an orthodoxly Kantian or Husserlian transcendental ego (even though, as I somewhat controversially suggested, a touch of transcendentality can be maintained in Jamesian pragmatism, too). James's view is, as has often been noted, closer to the "neutral monism" favored by Bertrand Russell (at some points of his career, influenced by James) and Ernst Mach than to solipsism. It is also possible to interpret his metaphysics as a panexperientialist or panpsychist process philosophy along the lines later explored by well-known but somewhat marginalized thinkers such as A.N. Whitehead and Charles Hartshorne.[63] James's pragmatism, moreover, renders his conception of human experience more "world-involving" (and thus, in a sense, realistic) than is often acknowledged. His emphasis on practical – and inevitably embodied – human action in a concretely experienced world actually brings him closer to, say, Martin Heidegger's and Maurice Merleau-Ponty's phenomenology than to Husserl's or Jean-Paul Sartre's approach.[64]

---

[61] See Sprigge, *James and Bradley*, especially Ch. 2.

[62] *Ibid*.

[63] Sprigge (*ibid.*) sees panpsychism as the unifying core of the otherwise rival philosophies of James and Bradley. According to panexperientialism (one form of which is panpsychism), one's own experience exemplifies the basic stuff of reality to be found everywhere (see *ibid.*, 45–6, and especially 441ff.). Bradley's Absolute is a "single vast cosmic experience" (*ibid.*, 265), the totality of the experience the world consists of, but Sprigge argues that Bradley's views should not be construed solipsistically any more than James's (*ibid.*, 367, 530). While panpsychism is different from solipsism, James himself understood that it might reduce the entire phenomenal world "to a set of solipsisms," thus perceiving the danger that his radical empiricism might after all lead to something like (multiple) solipsism (cf. William James, *Manuscript Essays and Notes*, in Frederick Burkhardt, Fredson Bowers, and Ignas K. Skrupskelis (eds.), *The Works of William James*, 19 vols. [Cambridge, MA and London: Harvard University Press, 1986], 107). Panpsychism and solipsism must, then, be distinguished, but this does not mean that there is no connection between the two. (In addition to Sprigge's work, see, for another recent defense of – primarily Whiteheadian – process-philosophical panexperientialism, David Ray Griffin, *Religion and Scientific Naturalism: Overcoming the Conflicts* [Albany: SUNY Press, 2000]).

[64] For some comparisons between these philosophers' treatments of the solipsism issue, with further links with Wittgenstein's relevant views, see Sami Pihlström, *Solipsism: History, Critique, and Relevance* (Tampere: Tampere University Press, 2004). A further comparison might be made between James's pragmatism and the American tradition of *personalism*, developed by "Boston personalists" like Borden P. Bowne and Edgar S. Brightman, because it is the concrete, living and acting human person that is James's focus in his explorations of experience (again nicely illustrated by his treatment of religious experience, in particular). The person, the subject of (religious) experience, is according to both James and the personalists irreducible, not to be accommodated by a reductive scientific (materialistic) world-view, or by the absolute of the Hegelian monists and pantheists. See Sami Pihlström, "Pragmatism and American Personalism: On the Possibility of Perspectival Metaphysics," *Iyyun* 53 (2004), 287–324. I shall not, however, further elaborate on these interpretive possibilities here.

Whether or not such comparisons turn out to be illuminating, avoiding solipsism was, it seems, a personal, agonizing – existential – need for James himself throughout his career. Charlene Haddock Seigfried has convincingly argued that one of the key aims of James's entire pragmatist philosophy (and philosophical psychology) was to avoid solipsism. "The solipsism of the stream of consciousness," the fact that "my consciousness never partakes yours," was, according to Seigfried, a "life-long obstacle" for James, an obstacle "to resolving the issue of how we nonetheless share a common world."[65] We *do* share a common world, James was convinced. The problem is to account for this philosophically, given the metaphysics of radical empiricism – and the kind of transcendental reinterpretation of pragmatist philosophical psychology I have proposed. The point is, as Seigfried puts it, that the problem of solipsism will be settled, and can only be settled, with reference to our common human *praxis*.[66] We may even view this as a (qualified) transcendental argument, ending up with the pragmatist notion of an "always already" shared practice as an answer to the transcendental question of *how* a common world is possible (*given* that it indeed is, in spite of radical empiricism). According to Seigfried, James offers us a concrete hermeneutical analysis of our practice-embedded being in the world, an analysis which helps us to get rid of the threat of solipsism.[67] The "refutation" of solipsism can only be found in our "embodied action in the world" and the "phenomenal bodily communication" based thereupon.[68] What this actually means must, however, be spelled out in more detail.

One worry that immediately arises is, again, related to James's individualist bias: the pragmatic-existentialist subject for whom the rejection of solipsism might be taken to be an ethical, existential, "will to believe" commitment[69] is at bottom an individual human being, a world-constituting mind whose selective interests are, again, at work even in the question regarding solipsism and its rejection. Both Jamesian pragmatism and, say, Sartrean existentialism, despite all their differences, almost exclusively focus on such an individual subject. If we follow James in arguing against solipsism and for a more pluralist conception of experience and the experienced world, we must at some point, guided by our pragmatic interests, ourselves *choose* to be convinced by the argument and thus *decide* to be non-solipsists. Such an ethically loaded choice can, however, ultimately only be made by *me* as an individual subject, as an ethical quasi-solipsist.[70] Moreover, I must, each and every

---

[65] Seigfried, *William James's Radical Reconstruction of Philosophy*, 92.

[66] *Ibid*., 188.

[67] *Ibid*., 189ff.

[68] *Ibid*., 279. Here one may argue that James's views are, again, relatively close to the phenomenology of the body later developed by Merleau-Ponty. (At this point, I am again indebted to an exchange of ideas with Sara Heinämaa). On James's worries about the problems of solipsism and other minds, see also Michael H. DeArmey, "William James and the Problem of Other Minds," *The Southern Journal of Philosophy* 20 (1982), 325–36, and Cormier, *The Truth Is What Works*, 16–8.

[69] Cf. William James, *The Will to Believe and Other Essays in Popular Philosophy* (1897 [1979]), Ch. 1.

[70] See Pihlström, *Solipsism*, Ch. 5.

time, make my choice *now*, as a quasi-temporal-solipsist – since, as perhaps Sartre more forcefully than anyone else has taught us, we have to choose our lives again and again, all the time, always at *this* precise moment.[71] Our individual ways of giving up solipsism and favoring a pluralism of experiential points of view (which, if we follow James, are all to be found within pure experience) instead may, therefore, be incomprehensible to others, or even to our own past and future selves. The problem of solipsism might thus seem to reappear just when we thought we had left it behind – in our very rejection of solipsism itself, in the commitment that we are prepared to make as ethically concerned pragmatic-existential creatures "concretely" living amidst others. The normative criteria that lead us away from solipsism depend on our private approval. Perceiving this, we are back in square one, engaging in our personal, individual existential situations.[72]

What these scattered remarks on the possibility that the problem of solipsism might come back should lead us to is a serious consideration of the metaphilosophical relevance of the individuality of the (transcendental) interest-guided subject conceptualized in Jamesian pragmatist philosophical psychology. Solipsism, the apparent result of our taking seriously such individuality, may not only teach us something about our subjectivity and intersubjectivity, but also about what we are actually doing when we philosophize – about solipsism, about the mind, or about anything else. Philosophy, including philosophical psychology, is *my* business. It is ultimately up to me to decide what *I* take to be an adequate philosophical picture of the world *I* experience, even if that picture crucially contains anti-individualist idea(l)s. This seems to be something that James acknowledged in his famous discussion of "philosophical temperaments" as formative factors in the history of philosophy.[73] And since solipsism *could* be true, after all, only about me and for me, its falsity is also something that concerns me first and foremost, in my existential solitude, however strongly my temperament commits me to anti-solipsist views, such as tolerance toward other perspectives on the world.

The fact that the possibility of solipsism can never be philosophically overcome once and for all, or grounded in a universally acceptable demonstrative argument, is, however, only something to be expected. Jamesian pragmatism was never designed to provide us with any metaphysical guarantees. On the contrary, it is a philosophy urging us to live forward in what James saw (and encouraged his readers to see) as a highly insecure world of real adventure, of real personal sorrows and pains, and of real gains and successes, for that matter. Our commitment to a pluralistic, non-solipsistic picture of experience and of the objects of experience ought to be

---

[71] I shall not discuss Sartre's existentialism here; see Jean-Paul Sartre, *Being and Nothingness: A Phenomenological Essay on Ontology* (1943), trans. Hazel E. Barnes (New York: Philosophical Library, 1966). For an existentialist reading of James's "will to believe" doctrine, see, however, Hilary Putnam, *Renewing Philosophy* (Cambridge, MA and London: Harvard University Press, 1992), Ch. 9.

[72] For a more thoroughgoing discussion of the "return" of the solipsism issue, see Pihlström, *Solipsism*, Chs. 4–5.

[73] James, *Pragmatism*, Ch. 1; on the relevance of philosophical temperaments, see also Pihlström, *Pragmatism and Philosophical Anthropology*, Ch. 10.

based on a healthy fallibilist, anti-foundationalist pragmatism which does not seek to prove in advance that solipsism is false. No such proof is forthcoming, any more than a foundationalist, non-pragmatic proof of pragmatism itself can be expected. In this uncertain situation, it is all the more important to commit oneself ethically in the concrete circumstances of life, as James (along with other existentialist thinkers, like Sartre) urged us to do.

On the basis of James's key starting point, pragmatic individualism, it seems, then, that the solipsism issue is inevitable, even though James tried hard to avoid it. In the remainder of this chapter, we shall further see how the issue of solipsism is inevitable even within a pragmatic realism about other minds. Yet, the fact that a given problem inevitably arises as soon as one thinks about it carefully enough does not mean that there is no way of living responsibly and thoughtfully with that problem, even throughout one's entire life. James's philosophical psychology might be seen as an attempt to teach us a way to live with the uncertainty we cannot eliminate.

11.4 An Argument from Analogy – Or an Ethical Relation to Otherness?

James himself was not a solipsist – this much is clear. But we may question the successfulness of the considerations that he, when theorizing about radical empiricism, offered against solipsism; in particular, we may ask whether the argument he gave is, at bottom, an "argument from analogy" in favor of other minds. That is, even if the other person (or mind) *is* realistically postulated, and solipsism abandoned, this pragmatic realism about other minds may still remain subordinated to my privileged experiential point of view.

At this point, turning to Josiah Royce's criticism of James, as presented in one of Royce's major works, *The Problem of Christianity* (1913) may help us to further understand James's position and its difficulties. Royce (the "absolute pragmatist") is much closer to Peirce than to James, and he explicitly states his indebtedness to Peirce's views, especially regarding the concept of *interpretation*, which is highly central in Part II of *The Problem of Christianity*.[74] From the perspective of the present inquiry, it is interesting to note how Royce argues against James's excessive individualism. It is from this point of view that we need to take notice of Royce – as a more or less Peircean voice in the pragmatist tradition – and particularly of his attack on James.[75]

---

[74] See Josiah Royce, *The Problem of Christianity* (Washington, D.C.: The Catholic University Press of America, 2001; reprint from the University of Chicago Press edition, 1968 [1913]), especially Chs. 11–14, or 275ff.; see also the preface, 39.

[75] This is not to deny that Royce was also greatly influenced by James's views. For discussions of the relations between these two philosophers, see Perry, *The Thought and Character of William James*, and especially Sprigge, *James and Bradley* (the latter focuses on Bradley rather than Royce, though).

Royce says he owes a great deal to James's *Varieties*, but the form of religious experience he describes differs "very profoundly" from the one James described "in one very important respect": "He [James] deliberately confined himself to the religious experience of individuals. My main topic is a form of social religious experience, namely, that form which, in ideal, the Apostle Paul viewed as the experience of the Church. This social form of experience is that upon which loyalty depends."[76] It was James's "profound" and "momentous" error to believe that such a social religious experience must be shallow or insincere.[77] This contrast to James leads to the heart of Royce's position, to the religion of "loyalty", i.e., loyalty to, or love of, an ideal (interpretative) community, a view developed throughout his book. It is (only) the loyalty to what Royce labels the "Beloved Community" that can save the individual; this must be contrasted to an individual imitation of Christ, the founder of the religion.[78] Thus, the psychology of the origins of Christian experience is social, not individual.[79]

Royce was a Peircean also in arguing that reality ultimately depends on interpretation. The "real world" means "true interpretation" of a problematic situation; there is no real world unless both the interpreter and the appropriate community are real.[80] *"The world is the interpretation of the problem which it presents."*[81] In relation to this Peircean-like, socially oriented metaphysics, Royce accused James of relying on what he found a superficial *analogy argument* for the existence of other minds:

> "Why," says Professor James, addressing a supposed fellow-man in one of his essays on Radical Empiricism, "Why do I postulate your mind? Because I see your body acting in a certain way. Its gestures, facial movements, words, and conduct generally are 'expressive,' so I deem it actuated, as my own is, by an inner life like mine. This argument from analogy is my *reason*, whether an instinctive belief runs before it or not. But what is 'your body' here but a percept in *my* field? It is only as animating *that* object, *my* object, that I have any occasion to think of you at all."[82]

---

[76] Royce, *The Problem of Christianity*, 40.

[77] *Ibid.*, 41.

[78] *Ibid.*, 45 and passim.

[79] *Ibid.*, 224. See also John E. Smith, "Introduction" (1968), in Royce, *The Problem of Christianity*, 8–9 (on the contrast to James), 32–3 (on overcoming the "twin evils of individualism and collectivism"). This may be compared to Taylor's more recent criticism of James in *Varieties of Religion Today*.

[80] Royce, *The Problem of Christianity*, 337, 339.

[81] *Ibid.*, 361 (emphasis in the original).

[82] Royce is quoting from James's "The World of Pure Experience" (1904; Ch. 2 in *Essays in Radical Empiricism*), 77–8. James goes on: "If the body that you actuate be not the very body that I see there, but some duplicate body of your own with which that has nothing to do, we belong to different universes, you and I, and for me to speak of you is folly. Myriads of such universes even now may coexist, irrelevant to one another; my concern is solely with the universe with which my own life is connected. In that perceptual part of *my* universe which I call *your* body, your mind and my mind meet and may be called conterminous." (*Ibid.*, 78.) It seems that James fails to get rid of the subject/object dualism here. (One may also suggest that both James and Royce are too heavily focusing on sight, or visual experiences, here. This was pointed out by Simo Knuuttila in discussion. I am not sure, however, whether transforming the examples to other sensory modalities would affect the argument in any significant manner).

> In the form of this familiar argument from analogy, – an argument which many philoso-
> phers indeed regard as expressing our principal reason for believing that our neighbors'
> minds are realities, – James also puts into words an equally familiar aspect of the meta-
> physical view which naturally accompanies this [Schopenhauerian] affirmation of the will
> to live. I perceive my own inner life, or, at all events, my own facts of perception. By analogy
> I extend the world thus primarily known to me. Other men are, in this way, hypothetical ex-
> tensions of myself. [. . .] Pragmatism, in its recent forms, is indeed one of the most effective
> philosophical expressions which Schopenhauer's "Will to live" has ever received.[83]

Thus, in effect, though not using this word, Royce accuses James of arriving at, or relying on, a kind of methodological solipsism, if not full-blown solipsism. As he further elaborates:

> For James [. . .], my only and, to his mind, my sufficient ground for believing in my fellow's
> existence, for "postulating your mind," is an argument from analogy, – an extension of the
> inner life of my already known self, with its feelings, with its will, and with the workings
> of its ideas, into the perceived body of my neighbor, whose movements and expressions
> resemble mine.
>   Now, as a fact, the most important part of my knowledge about myself is based upon
> knowledge that I have derived from the community to which I belong. In particular, my
> knowledge about the socially expressive movements of my own organism is largely derived
> from what I learn through the testimony of my fellow-men. Therefore I cannot use the
> analogy of our externally expressive movements as my principal reason for believing in the
> reality of the inner life of my fellow-man, because I am very largely unable to perceive my
> own expressive movements in as direct a way as is that in which I perceive the organism
> and the movements of my fellow-man.[84]

That is to say, my "postulating" other minds cannot be grounded in the analogy, because the workability of the analogy already presupposes the community in which I live and in which the reality of other minds is not doubted.[85] Royce further asks: "Do I reason thus: 'When my face looks thus, I feel so and so; therefore, since my neighbor's face looks thus, it is fair to reason by analogy that he feels so and so?' How utterly foreign to our social common sense would be this particular argument from analogy!"[86] To "postulate your mind" on the basis of the analogy, he says, would be "a form of solemn fooling". My belief in my fellow's mind is *prior to* my belief in the expressibility of my own gestures; the latter cannot be used as a piece of evidence for the former, *contra* James.[87]

It is on these grounds that Royce arrives at his own (in his view clearly anti-Jamesian) conclusions: "I postulate your mind, first, because, when you address me, by word or by gesture, you arouse in me ideas which, by virtue of their contrast with my ideas, and by virtue of their novelty and their unexpectedness, I know to be

---

[83] Royce, *The Problem of Christianity*, 353.

[84] *Ibid.*, 358.

[85] Is this a transcendental argument, i.e., an argument proceeding from something that is taken as given to a conclusion regarding its conditions of possibility? I leave this question for my readers to think about.

[86] *Ibid.*, 359.

[87] *Ibid.*, 359–60.

not any ideas of my own."[88] And further: "The reason, then, for 'postulating your mind' is that the ideas which your words and movements have aroused within me are not my own ideas, and cannot be interpreted in terms of my own ideas, while I actually hold, as the fundamental hypothesis of my social consciousness, that all contrasts of ideas have a real interpretation and are interpreted."[89] Thus, without entering any details of Royce's theory of loyalty and the "Beloved Community", or of his conception of "Pauline" Christianity based upon this notion, it is instructive to contrast his social/communal form of pragmatism (which is, to a large extent, derived from Peirce's similarly oriented theory of interpretation) to James's individualistic philosophy and psychology (of religion).

Still, there may be more in James than meets the eye.[90] In particular, James (or at least the Jamesian pragmatist who is willing to reconstruct rather than merely interpret James's views) may hold, or may be interpreted to hold, ultimately, that our relation to other human beings is *not* primarily a cognitive relation at all. Despite his "evidentialist"-sounding,[91] analogy-based treatment of the other minds problem, we may also read James as affirming the fundamentally *ethical* quality of our relations to the fellow humans surrounding us.[92] This would bring him closer to the Wittgensteinian or even Levinasian view according to which an "attitude towards a soul", an ethical attitude to the other person, is prior to any theoretical metaphysical or epistemological doctrine about there being an entity called the soul, or the mind,

---

[88] *Ibid.*, 360.

[89] *Ibid.*, 361.

[90] Even though I am about to defend James here, this is not in order to diminish Royce's importance (or the importance of other socially oriented religious thinkers, e.g., Taylor). I merely claim that James can be partly rescued from the excessive individualism these critics find in his thought (cf. also Pihlström, "Mortality, Individuality, and Pluralism") – although, as will become clear, I think James *cannot* be rescued from the "return" of the problem of solipsism as a transcendental problem. For another argument to the effect that the classical pragmatists, James included, did not infer the reality of others on the basis of an analogy, see Rosenthal, *Speculative Pragmatism*, 87: "[O]ur experience of other ontologically thick persons, whose commonness and uniqueness are manifest in their creative interaction, is a primary fact of experience. We do not begin with subjective contents as constitutive of our own person, and then from the outward behavior of another being 'infer' that this, 'like me,' is a person [...]. The ontological depth and density of other persons with whom I interact is something I experience; I do not infer to them [...]."

[91] I am here using the word "evidentialism" analogously to the way in which it is used in modern philosophy of religion, viz., in this case, roughly referring to any view which says that the existence of other minds is an hypothesis requiring supporting (or falsifying) evidence.

[92] Alternatively, given that the meaning of "cognitive" is pragmatically widened in James's will to believe considerations (see James, *The Will to Believe*), we may say that the "postulation" of other minds is cognitively acceptable on the basis of a will to believe "leap", i.e., in the absence of conclusive evidence but in a situation in which such a belief is a genuine need in our doxastic and especially ethical lives. The will to believe strategy, as I have argued elsewhere (Pihlström, *Pragmatism and Philosophical Anthropology*, Ch. 6), thus has a much broader area of application than philosophy of religion. For comparisons between the will to believe argument and the project of James's *Varieties*, see Proudfoot (ed.), *William James and a Science of Religions*; for a recent analysis of the will to believe argument, cf. also Gale, *The Philosophy of William James*, especially Ch. 4.

that the other possesses – or about our knowing, or being justified in believing, that there is such an entity.[93]

What I have in mind in suggesting this re-reading of James is primarily his critical discussion of our instinctive "blindness" to other humans' goals and projects, a blindness we should continuously struggle to overcome.[94] Ethical interests and needs are centrally involved among the ones directing our selective attention, thus crucially shaping our pragmatic engagement with – and our transcendental constitution of – the world we live in. I do admit, though, that these discussions may also be interpreted as instances of the analogy argument. I am, James seems to be telling me, instinctively blind to the fact that others posses an inner life and purposive (selective) interests just as I do. But there are elements in James that may support the view that, for him, the problems of solipsism and of other minds were not *purely* metaphysical and/or epistemological but primarily ethical. Thus, I suggest, we may (re-)read him – in what might be described as the post-Wittgensteinian situation in ethical and metaethical thought – as a philosopher who favored an ethical attitude toward a soul as a primary relation to other people as compared to any metaphysico-epistemological demonstrations of the reality of other souls.[95] We might, that is, rearticulate the experiential field of James's radical empiricism (that is, what he called pure experience) as an always already ethically structured field, something whose metaphysical constitution cannot be described in ethically neutral terms at all.[96]

We may also say that, here and elsewhere, the Jamesian pragmatist may infer a metaphysical conclusion, in this case, the commitment to the reality of minds, especially "other minds", from an ethical premise, in this case, our – or, primarily,

---

[93] The contrast between the ethical and the cognitive, as employed here, signifies no commitment to metaethical non-cognitivism but only an indication that others are ethically acknowledged rather than (propositionally) known to exist, let alone known to exist on the grounds of a metaphysical theory postulating substantial souls or other mysterious (material or immaterial) entities.

[94] See James's 1898 essay, "On a Certain Blindness in Human Beings," reprinted in William James, *Essays in Religion and Morality* (1982). I discuss this topic further in Pihlström, "William James on Death, Mortality, and Immortality". Cf. also James's "democratic" conception of immortality formulated in another 1898 paper, "Human Immortality," also in James, *Essays in Religion and Morality*. Gary Hatfield suggested (in conversation) that we might read James as saying (in the passage Royce quotes, and in related ones) that I may directly perceive qualitative (and possibly ethical) states like happiness as expressed on the other's face. This fits my reading very well. The ethical relation to the other surely has a perceptual component.

[95] See Wittgenstein, *Philosophical Investigations*, II, iv. Cf. here my reading of such Wittgensteinian philosophers' as Raimond Gaita's, Rush Rhees's, and Peter Winch's views on ethics in Sami Pihlström, *Pragmatic Moral Realism: A Transcendental Defense* (New York and Amsterdam: Rodopi, 2005). These authors must be neglected here; see, for instance, Gaita's *Good and Evil: An Absolute Conception* (2nd ed., London and New York: Routledge, 2004 [first published in 1991]) for an inspiring account of the "mysteriousness" of other human beings and their distinctive points of view on the world. (None of these Wittgensteinians seems to have much interest in James, unfortunately).

[96] For a defense of pragmatic moral realism by means of transcendental argumentation, see Pihlström, *Naturalizing the Transcendental*, Ch. 7, and *Pragmatic Moral Realism*.

my – concretely experienced need to take the other into account ethically. Such a step from ethics to metaphysics is, it seems to me, fundamental in James's pragmatism, although it is (once again) perhaps safer to regard it as a rational reconstruction of his views rather than a historically faithful interpretation of what he really thought. A non-pragmatist can hardly take such a step, but for someone already committed to pragmatism it is a natural one to take; moreover, among one's pragmatic reasons for committing oneself to pragmatism may indeed be the possibility of such an ethical grounding of metaphysical ideas. What this amounts to in the case we have considered is, in brief, the acknowledgment of other people, "other minds", or, preferably, other human persons[97] in the absence of any prior metaphysical and/or epistemological argument or "proof" that would justify such an acknowledgment. This conclusion is independent of whether such an argument would be an argument from analogy, as discussed above, or a different kind of argument – or even some non-argumentative metaphysical "ground" for the postulation of others, such as the "empathetic intuition" and "I-Thou" relation to others that James's mysticism invokes.[98] Our acknowledging the experiential perspectives of others is conceptually prior to any such allegedly justificatory reasoning.[99] The other's perspective is, indeed, ethically demanding without any prior justification of its reality; argumentation comes after the fundamental but ungrounded demand of justice.

This is the sense in which I would like to reinterpret the initially puzzling "mind-dependence" of the ontology of the mental realm. Postulating minds is grounded in our experiencing others as ethically valuable creatures, ideally in a concrete face-to-face relation to them. In *this* sense, ethics is, for James as much as for Emmanuel Levinas, a "first philosophy".[100] Ethics is a transcendental ground

---

[97] On the distinction between minds and persons, and the corresponding distinction between *mindism* and *personism*, see Elmer Sprague, *Persons and Their Minds: A Philosophical Inquiry* (Boulder, CO: Westview Press, 1999). In these terms, James would undoubtedly qualify as a personist: for him, as we have seen, the mind is not a soul-like immaterial entity within a material body – or reducible to the material body – but an activity of a whole human self, of a person engaged in worldly practices.

[98] See, again, Gale, *The Philosophy of William James*, especially Ch. 9. Ultimately, according to Gale, James the mystic wanted to "go all the way" toward an "I-Thou" relation to the entire (personalized) universe, as nature mystics had done (*ibid.*, 183).

[99] The same point, *mutatis mutandis*, goes through, I believe, when Jamesian pragmatism is applied in the philosophy of religion and philosophical thanatology (see my essays on these issues cited above), as well as in moral philosophy more generally. Let me note that the notion of acknowledgment is adopted from Stanley Cavell's writings, which I cannot comment upon here, though (cf. again Pihlström, *Solipsism*, Chs. 4–5).

[100] Instead of dealing with Levinas's thought in any detail, I simply refer the reader to his central writings collected in Sean Hand (ed.), *The Levinas Reader* (Oxford: Blackwell, 1989). Comparisons between Levinas and pragmatism have up to now been rare; a noteworthy exception is Hilary Putnam, who both subscribes to a Jamesian-Deweyan form of (neo)pragmatism and takes Levinas seriously: see Putnam, "Levinas and Judaism," in Simon Critchley and Robert Bernasconi (eds.), *The Cambridge Companion to Levinas* (Cambridge: Cambridge University Press, 2002), as well as Putnam, *Ethics without Ontology* (Cambridge, MA and London: Harvard University Press, 2004).

(itself ungrounded) for any metaphysics (of the mind or anything else), because it provides, pragmatically, a key source of the human interests that are relied upon in any humanly possible categorization of reality (minds included). To be sure, James, together with a number of other pragmatists, often seems to assume a consequentialist and utilitarian conception of ethics; yet, his moral philosophy is hardly reducible to a simple utilitarianism. If it were, a comparison to Levinas or to Wittgenstein would hardly be plausible. A commitment to a fundamental ethical perspective on human life, and to ethical duties to other humans not to be accounted for in utilitarian terms, runs through James's philosophical writings and is clearly visible, e.g., in his 1891 essay, "The Moral Philosopher and the Moral Life".[101] Echoing the later Wittgensteinians, I would like to claim that ethics, for a Jamesian pragmatist, is *not* a special subject-matter or discipline of academic philosophy. After all, the "Moral Philosopher" article is his only work specifically on ethics. For James, ethics lies, rather, everywhere, in each and every corner of our experience.[102] Furthermore, the crucial pragmatic outcome that my belief in other minds, or the reality of other human persons, may yield is, precisely, my ethical ability to genuinely take the other's perspective into account in my actions. In the interest-guided selectivity of our mental activities, the ethical qualities of the world we construct should, then, be granted a primary place. It is precisely these ethical aspects of our being in the world with others that we should first and foremost direct our attention to. This "should" is itself in the ethical mode, precisely because no prior philosophical theory can justify or lay foundation to ethical relations between concrete individuals.

The account I have sketched is a picture of James that we may end up with insofar as we emphasize the central place of pragmatism – the search for (conceivably) pragmatically significant "results" of one's beliefs – in James's thought. It is only if we place his radical empiricism before or above his pragmatism (and sharply separate between the two) that we are led to the problems of solipsism and other

---

[101] This essay is reprinted in James, *The Will to Believe*. Gale also refers to James's "deontological intuitions" and notes that he often "writes like a good Kantian" (*The Philosophy of William James*, 35).

[102] At this point, the Jamesian pragmatist may come close to the picture of the fact/value entanglement defended by Hilary Putnam in some of his recent writings: see Putnam, *The Collapse of the Fact/Value Dichotomy and Other Essays* (Cambridge, MA and London: Harvard University Press, 2002). In his paper, "Philosophy as a Reconstructive Activity: William James on Moral Philosophy," in William Eggington and Mike Sandbothe (eds.), *The Pragmatic Turn in Philosophy: Contemporary Engagements between Analytic and Continental Thought* (Albany: SUNY Press, 2004), 31–46, Putnam draws our attention to James's "The Moral Philosopher and the Moral Life," in a manner that seems congenial to the "ethics first" line of interpretation I am proposing here. James maintains, according to Putnam (*ibid.*, 35), that "all obligation [. . .] and all standards [. . .] can only arise from the demands of sentient beings. When something is 'required by us', it *is* literally required ('demanded' is James' word for this) by some sentient being or other, and it is, so to speak, the will of the other that we experience as a demand upon us, and nothing more than this." Although I want to acknowledge my debt to Putnam's perceptive readings of James, it should be noted that Putnam would be as unhappy as James himself with the vocabulary of the "transcendental" that I have employed.

minds, and to the argument from analogy whose credentials Royce rightly treats skeptically. Insofar as pragmatism, instead of the metaphysics of pure experience (or any other philosophico-psychological metaphysical theory) is taken to be the true core of James's philosophy (and philosophical psychology),[103] we may say that the problems of solipsism and of other minds can be resolved, or at least we can go on living with these problems. What I mean is that we may take it as an ethically grounded (and, therefore, *not* metaphysically or epistemologically grounded) aspect of human life as we know it that we simply do take others into consideration in our thought and action. There is no more ultimate, non-pragmatic foundation for this attitude to be found anywhere within the realm of human experience. The ethical attitude, moreover, is constitutive of our practices, and hence of our ontological, constitutive structuring of reality (including mentality) through those practices. Thus, the way(s) in which the world is cannot be metaphysically determined independently of our ethical perspectives – which is one way of stating the Jamesian thesis of the dependence of ontological categories on purposive human action.

Seigfried succeeds in expressing the ethical motivation – the profound ethical seriousness – of James's philosophico-psychological reflections in a thoughtful manner:

> Our interest in the mind/body problem is not an idle one of merely gathering facts for their own sake, as the positivist attitude of disinterested scientific neutrality would have it. We want to know whether we are responsible for our activities or are determined by events outside of our knowledge and control. [. . .] The issue is precisely whether events which we experience as ours are in fact so, or whether they should be reductively attributed to brain cells. [. . .] The only reason for investigating activity and causality is to help us understand the course and meaning of life. The pragmatic stance is that we seek to know, not for its own sake, but to enable us to live better.[104]

This formulation points toward James's life-long interest in the problem of determinism vs. indeterminism, which is of course a metaphysical problem but which above all had a crucial ethical relevance by James's lights. Indeed, nothing in philosophical thinking, or in *any* humanly relevant thinking, can be "disinterested", according to James. We are always already engaged in our interest-driven practices. In particular, when we are doing (philosophical) psychology or philosophizing about the mind, we ultimately attempt to find ways to live better. In this sense, James's philosophical psychology attempts to keep one of the ancient promises of philosophy, to put our understanding of our mental activities (the Socratic *gnōthi seauton* project) in the service of the good life, thereby also self-critically evaluating the very meaning of "the good life".

---

[103] Again, I shall not try to determine what James himself actually thought about the relation between pragmatism and radical empiricism. Occasionally, he seems to think these two were independent doctrines; on the other hand, pragmatism seems to require an appeal to the ubiquity of experience that a metaphysics of pure experience as the ultimate stuff of reality helps to sustain. If the points I have made are on the right track, the metaphysics of radical empiricism cannot get off the ground unless a prior pragmatic – ethical – commitment to the significance of experience other than mine has been made.

[104] Seigfried, *William James's Radical Reconstruction of Philosophy*, 322–23.

11.5 Conclusion

This investigation has, I hope, provided one example of the way in which problems in metaphysics and epistemology – and, correspondingly, in the philosophy of mind or philosophical psychology – are subordinated to ethics, or at least should be, insofar as we follow the basic insights of Jamesian pragmatism. It should be clear that I have not attempted to argue for pragmatism in general terms at all. The results of my investigations should be formulated in a conditional manner: *if* Jamesian pragmatism is adopted as one's basic philosophical orientation, *then* one will be justified in treating the issues of the mind and solipsism in the ethically grounded way I have outlined. Whether or not this constitutes an indirect pragmatic argument for pragmatism depends on one's more general philosophical assumptions. I am inclined to think it does constitute such an argument, but I understand that not all philosophers would share my basic intuitions here.

The Jamesian pragmatist may, I conclude, subscribe to the following transcendental picture. The world, including in particular the reality of the human mind(s) there are in the (empirical) world, gets structured through our interest-driven practices, through vital human needs and their satisfaction, and especially – I have argued – through our ethical need to take into account other (potentially richer) perspectives on the world than our own. Again, we may emphasize the central pragmatist point that it is only in a context defined by our pragmatic, particularly ethical orientation in the world that the world itself, including the mind(s) there are, receives any ontological structure; there is no mental, or other, reality independently of our pragmatic-cum-ethical need to postulate such a reality. Practices in general, as I urged above, play a transcendental role; ethical practices, or practices of moral deliberation, in particular, play such a role in the constitution of the ontology of the human world.[105]

Moreover, if these reflections are correct, then radical empiricism will require substantial modification. The pragmatic and especially ethical needs and interests that ground our postulation of mentality are above all *mine*; they are not just neutral bits of pure experience. If the reality of other persons is "carved up" or constituted by our ethical needs and interests, this "carving up" must be done, transcendentally, by me, the one and only ethically engaged subject, the one for whom ethics is (or should be, at least) a fundamentally serious concern. In true pragmatism, there is no way to deny subjectivity in favor of an all-inclusive "experience". It is in me that I feel my ethical obligation to others (or, again, should do so); it is in and through my concrete phenomenological encounter of the other that ethical duty arises – for both James and a thinker such as Levinas. Here, no argument from analogy functions;

---

[105] Insofar as the ethical acknowledgment of the other person and her/his perspective on the world is fundamental in James's pragmatism, the Jamesian approach to subjectivity and experience might even amount to an investigation of "transcendental intersubjectivity" in a sense comparable to the phenomenological investigations of intersubjectivity in Husserl's philosophy (cf. again Dan Zahavi's writings cited above).

I am irreplaceable and unique.[106] Yet, it is – uniquely – my responsibility *not* to acknowledge otherness simply on the basis of analogies to my own case. I can see no way out of this tension, but I do not consider this situation disastrous. On the contrary, the tension I hope to have identified illuminates the foundationlessness and irreducibility of ethical duty.

We may note, finally, that it is through the *transcendentality* of James's approach that the true significance of the intimate connection between metaphysics and ethics receives illumination. The pragmatic acknowledgment of other human persons and their experiential perspectives is a condition for the possibility of any tenable metaphysical position, because metaphysical commitments require human practices, habits of action, in which certain interests and purposes are, though endlessly debatable and reconceptualizable, at least temporarily accepted and strived after. Such interests and purposes can never be individually established, let alone achieved, because we find ourselves within a field of human interests not all of which are our own. To this extent we may accept Royce's claim that the argument from analogy in favor of other minds cannot even get started unless we already assume that we live in an interpretative community of fellow humans, even though Royce's criticism of James may have been too harsh. (We may say that *if* James offered an analogy argument, as Royce claims, then his position ought to be rejected.) Indirectly, we may also arrive at a metaphilosophical point about the admissibility of transcendental reasoning – and even transcendental psychology – in pragmatism: establishing the transcendental conditions for the possibility of experience in the form we find it in our lives requires ethical orientation, in which other experiential perspectives are treated with respect. It is in our concrete encounter with such other perspectives – other human beings – that we may locate some of the necessary conditions for the possibility of structured experience. But, again, we should not overlook the fact that this kind of ethically engaged pragmatism does (as I have argued) in a sense place subjectivity – *my* subjectivity – first, requiring that I take the responsibility for the other, avoiding the "blindness" that James was worried about. In consequence, pragmatism never entirely gets rid of the problem of solipsism; nor does, I believe, any other truly transcendental philosophy.

Let me repeat that whether or not James's own pragmatism and radical empiricism ultimately avoided solipsism (and, if they did, how) is beside the point here. It is, rather, both philosophically and historically important to note that the tradition of pragmatism, even in its allegedly radically individualistic Jamesian form, does

---

[106] For a discussion of how this leads to a re-emergence of solipsism on the ethical level (with some more detailed remarks on Levinas's ethics of otherness), see again Pihlström, *Solipsism*, Ch. 5. The point is that there is a sense in which (transcendental, ethical) solipsism is right even when connected with a (pragmatic, empirical) realism about other minds: ethical obligation, in order to be ethical at all, must be conceived as uniquely *mine*. Insofar as this is a fair re-reading of James's picture of the mind and other minds, one of my conclusions is that the Jamesian pragmatist cannot avoid the solipsism issue.

possess resources to deal with these fundamental issues in the philosophy of mind in an ethically responsible manner. And it is at least as important to note that these issues may, even for a pragmatist, nevertheless remain unresolved on a transcendental level.

Chapter 12
Psychology and Metaphysics from Maine de Biran to Bergson

Pascal Engel

Abstract The chapter examines the complex relationship between psychology and philosophy as it was conceived by the French "spiritualist" tradition of the nineteenth century. Philosophers like Maine de Biran, Lachelier, and Bergson in particular hoped to move from a psychological (mostly introspective) analysis of the human mind to metaphysical theses about the self, being and time. The views of Maine de Biran on trying and willed movement and those of Bergson on "durée" in particular are examined. It is suggested that the price paid by this speculative method in psychology is a complete divorce from actual psychological research, and an illusion of understanding.

Keywords French spiritualism · introspection · metaphysics · Maine de Biran · Lachelier · Bergson

What can psychology, understood as a study of the nature of mental phenomena, tell us about metaphysics, understood as an account of the nature of being and of the most general properties of the world? At a certain degree of generality, this question is itself a metaphysical one: how can our knowledge about our own minds tell us something about the world? But the meaning of the question depends in turn upon what one means by "knowledge about our own minds." If this knowledge is subjective and first person, then we get the familiar idealist question concerning the relationship between the contents of our minds and the nature of an objective external reality. Then again, if knowledge of our minds is objective, or can claim some form of objectivity, then the question becomes how the account of what humans think, feel, and do impinges upon our conception of reality. Psychology, along with sociology, history, and anthropology, belongs among the disciplines which deal with the latter question. There are many different ways of understanding the nature of this enterprise and these depend upon what one takes psychological properties to

P. Engel
Department of Philosophy, University of Geneva

involve. Correlatively, there are various ways of understanding our initial question concerning what psychology can tell us about metaphysics. Two of these are prominent.

One distinctive project proceeds along the lines of Hume's naturalism. It consists in the attempt to understand the origins of our "ideas", "representations", "concepts", starting from a study of "human nature" – which Hume himself conceived as a study of the laws of association of ideas. These "ideas", or "representations" or "concepts", concern objects, properties, existence, space, time, causality, necessity and possibility, and other such issues usually considered as pertaining to metaphysics or ontology. Such metaphysical notions are found to have a natural origin in our minds. A contemporary version of this naturalistic enterprise, which replaces associationist psychology with modern day cognitive science, consists in an investigation into our "naïve physics", our "natural theories of mind", or our common sense notions of object, property, or event. Alvin Goldman, for instance, explicates a project of this sort, and it is today pursued in various ways elsewhere.[1] This project purports to give us a *psychological* account of metaphysics. It is important to understand how this kind of descriptive metaphysics differs from the two traditional forms of metaphysics: speculative metaphysics and revisionary metaphysics, respectively. It differs from speculative metaphysics first in that it does not attempt to derive metaphysical theses from *a priori* reasonings or intuitions in the traditional fashion. It differs from revisionary metaphysics in that it does not try to replace our ordinary conceptual scheme by a better one as, for instance, when certain metaphysicians urge us to replace our tri-dimensional view of objects as continuants by a quadri-dimensional picture which identifies objects with temporal parts. Such a descriptive project gives us a sense of how psychology and other sciences can tell us something about metaphysical issues. The problem, however, is that this kind of project seems to toll the death knell of metaphysics instead of enriching it. At least within the Humean framework, an inquiry into the nature of the ideas of cause or necessity shows that these concepts are psychological and, moreover, nothing but psychological. The mind projects, or spreads itself onto the world, but the causes and the necessities are not in the world, only in the mind. The descriptive project seems to amount to the dissolution of metaphysics into psychology.

There is another way of understanding the relationships between psychology and metaphysics which is less deadly for the metaphysical enterprise but is bolder. It consists in an attempt to derive conclusions about metaphysical concepts, such as 'substance', 'self', 'causality', or 'necessity', from an introspective investigation into the contents of our own minds. This enterprise has an obvious idealist ring, and it raises the immediate question: How can an investigation into our purely subjective world be a guide in our attempt to understand and describe the nature of the objective

---

[1] See Alvin Goldman, *Philosophical Applications of Cognitive Science* (Colorado: Westview Press, 1998). This particular project has affinities with Strawson's descriptive metaphysics, although he did not mean to describe our conceptual scheme from a material taken from psychology, but from our common understanding of our language, see Peter F. Strawson, *Individuals* (London: Methuen, 1959).

world and the ultimate properties of reality? In many ways this question has been asked since the first use of the notion of 'psychology' in a metaphysical setting. This critical question can be directed at Descartes' attempt to argue for the real distinction of mind and body from the *cogito*. It can also be directed at Wolff's *psychologia rationalis*, which attempts to deduce necessary properties of the self as a substance from an investigation into the contents of our concepts. At this latter project, Kant strikes a deadly blow in the *paralogisms of pure reason*, by which he demonstrates that the very attempt necessarily fails.

In the nineteenth century, there was one school of thought that explicitly attempted to derive metaphysical theses from an investigation into the mind, namely, the school of French spiritualism. A number of French philosophers of the nineteenth century thought that they could proceed from an inner description of the mind to the metaphysical properties of the self and to ontology in general. The extent to which this kind of project can be pursued depends on the nature of the psychological investigation and its scientific character. The project of the French spiritualists is based on introspective psychology, and it was essential to the spiritualists that the contents of the mind be discovered from the first person perspective, even though they were supposed to be objective in some sense of the word. But spiritualism soon had to confront the results of an emerging science of psychology.

Here I would like to review some doctrines and methods of the spiritualist school, from Francois Marie Pierre Gontier de Biran to Henri Bergson, and to suggest why this bold attempt to disguise metaphysics in psychological clothing fails.

12.1 Maine de Biran: From the Bodily Self to the Substantial Self

In order to understand the prehistory of the thesis that psychology can guide metaphysics, it is beneficial to return to a classical discussion that originated in Descartes and his argument for the real distinction of mind and body. Descartes famously argues that since he has a clear and distinct idea of the main attribute of a thinking thing, and since he can conceive of this thinking thing as existing without the body, it follows that there is a real distinction between the mind (thinking substance) and the body (extended substance). Thus, he moves from what we *conceive* of the mind to what it *is*. But this move seems fallacious, as Father Bourdin remarks, *ab nosse ad esse non valet consequentia*: from the fact that we can conceive of mind and body as distinct it does not follow that they are distinct in reality. Descartes refuses this maxim and argues that when our thoughts and conceptions are clear and distinct and when we actually know the thing as it is we capture the real nature of the thing.[2]

---

[2] Descartes, *Responses to the Seventh Set of Objections*, in René Descartes, *Œuvres de Descartes*, vol. 7, ed. Charles Adam and Paul Tannery (Paris: J. Vrin, 1964–76), 519 (hereafter referred to as AT, followed by volume number). On Descartes' reasoning, see Bernard Williams, *Descartes: The Project of Pure Enquiry* (London: Penguin, 1978) and Tamar S. Gendler & John Hawthorne (eds.), *Conceivability and Possibility* (Oxford: Oxford University Press, 2002), 13–26.

Descartes has no place in his system for a separate science of the mind: on the contrary, he takes mind and body not only as separate but also as tied together by a substantial union. The *cogito* allows us to know the nature of the soul, but it cannot be the basis of a scientific study of the mind. As regards the objective study of the mind, the proper object is more specifically the body and not the inner realm of a thinking subject.

The view that psychology could be a science of the "inner sense" is not Descartes'. This perspective emerged during the eighteenth century with the British empiricists. In France, it was introduced by Etienne Condillac and more extensively, at the time of the French revolution, by the "Idéologues" (Destutt de Tracy, Jouffroy, Royer-Collard) who conceived the project of a purely introspective psychology. Maine de Biran (1766–1824) was their disciple, and he proposed that psychology could be based on the foundation which Descartes thought to be impossible:

> Whereas Descartes thought that he had put forward the first principle of all science, the first self-evident truth, by saying: I *think*, therefore I *am* (a *thinking* thing or substance), we would say better, in a more determinate manner, and this time with the undeniable self-evidence of the inner sense: I act, I want or I think the action, hence I feel myself as a *cause*, hence I am or exist really as a cause or a force. It is exactly under this relationship that my inner thought is the expression or the conception or the production of my own *real* existence and, at the same time, the primary manifestation and birth of the *self*, which is born for itself by starting to know itself.[3]

What Biran calls "the primitive fact" is a psychological fact in the sense of a being which is only accessible through introspection. This is the basis of all knowledge, subjective knowledge as well as objective knowledge. Hence, psychology is the basic science. But these primitive facts, and in the same way all the fundamental facts of psychology, are not merely psychological, in the sense of being mind-dependent and mind-relative. They are also actually *objective* facts. But we should not understand them in the same way as the facts of physics:

> There are two order of phenomena, distinct and even opposed, and consequently two sorts of observation, which have nothing in common in their means, their object, or their aim – and even seem generally to be at odds with each other. The one takes wing far from ourselves, the other as close as possible to the self, seeking only to incline in its depths.
>
> It is no doubt with this divergence in the means and the direction of the two sciences in mind that Newton, touching on the question that concerns us, cried "Physics, preserve me from metaphysics."
>
> We also, having in mind the observation, necessarily twofold, of two classes of phenomena of which the mixture and confusion offer so many errors, illusions and miscalculations, may cry out in turn "Psychology, preserve yourself from physics."[4]

---

[3] Francois Marie Maine de Biran, *Nouveaux essais d'anthropologie. Note sur l'idée d'existence. Derniers fragments* (1824), in *Œuvres*, ed. Bernard Baertschi. (Paris: J. Vrin 1989), 5 X–2.77. All translations are mine when not otherwise indicated. On Biran's thought, see especially Francis Charles Timothy Moore, *The Psychology of Maine de Biran* (Oxford: Clarendon Press, 1970); Bernard Baertschi, *Les rapports de l'âme et du corps: Descartes, Diderot et Maine de Biran* (Paris: J. Vrin, 1992); François Azouvi, *Maine de Biran: la science de l'homme* (Paris: Vrin, 1995).

[4] Francois Marie Maine de Biran, *Rapports du physique et du moral chez l'homme*, in *Œuvres*, ed. F.C.T. Moore (Paris: Vrin, 1984), 8.126.

The question then becomes: What is the nature of these primitive psychological facts? Maine de Biran's characteristic doctrine is that it is not cognition or thinking in general but the perception of our action or volition. The real "cogito", then, is not a cogitation but the feeling of what Maine de Biran calls "*effort*", willed movement, or the experience of trying.

The description of this experience of effort and willed movement is one of Maine de Biran's most important contributions. The body which is experienced in the attempt of trying to do something, such as raising one's arm, is not physical, in the sense of being external to us and extensive. On the contrary, it is constituted by *kinesthesis*; it is a bodily awareness, a lived body or "*corps propre*", to use the phrase that Maurice Merleau-Ponty has later made well-known. The lived body is thus distinct from the materiel body: the first is the object of our immediate awareness, whereas the second is organic. Maine de Biran's analyses of the bodily self had a deep impact on later French philosophy, on Bergson, Merleau-Ponty, and the French existentialists.[5]

Maine de Biran, however, refuses to directly infer the nature and existence of a substantial self from the mind's feeling of itself in action and volition:

> If I do not get out of the conscious fact, I could not find any support for this proposition: *I am a thinking being*, for my thought would have to be felt or perceived as the fundamental mode or permanent *attribute* of the substance, such that there existed in consciousness a real duality, or a relationship between two distinct terms, one of which being the substance and the other the mode or the attribute.[6]

Unlike Descartes, Maine de Biran does not accept the step from the mode of thinking to the thinking substance. The idea of the self does not contain the idea of the soul. In this sense, Maine de Biran is a phenomenalist: nothing exists in consciousness except the feeling of a relationship. The felt relationship is that of causality, which is experienced in action and bodily movement. It follows that the self only exists, according to Biran, at the moment in which it is perceived in effort and action. When we are asleep, for example, or are unconscious, the primitive fact is not there.

However, in a number of writings Maine de Biran rejects the consequences of a strict phenomenalist point of view and the "bundle conception" of the self associated with it. But neither does he want to assert the reality of the self as a perceived phenomenon. The self and the soul do not exist as reality. But they exist as *noumena*.

---

[5] See Maurice Merleau-Ponty, *L'union de l'âme et du corps chez Malebrache, Biran et Bergson* (Paris: Gallimard, 1968); Maurice Merleau-Ponty, *Phénoménologie de la perception* (Paris: Gallimard, 1945); Michel Henry, *Philosophie et phénoménologie du corps* (Paris: PUF, 1967). In contemporary anglophone philosophy, one writer whose views on action and willing have some affinities with Biran is Brian O' Saughnessy, *The Will* (Cambridge: Cambridge University Press, 1980), and to a certain extent the idea is rediscovered in contemporary philosophy of mind with the idea of a bodily self. See in particular Jose Louis Bermudez, Anthony Marcel, and Naomi Eilan, *The Body and the Self* (Cambridge: MIT Press, 1995)

[6] François Marie Maine de Biran, *Essai sur les fondements de la psychologie* (1812), in *Œuvres*, ed. F.C.T. Moore (Paris: Vrin, 2001), 7.1.

Through our experience of action we know only effects, and there is a cause of these effects, although it is unknowable.

This means that Maine de Biran does not make any inferences from the psychology of our inner sense to the nature of a metaphysical self. But his disciples, such as Victor Cousin (1792–1867), and the philosophers whom he inspired, such as Felix Ravaisson (1813–1900), drew the inference from introspection to metaphysics. They maintained that through reflection and the experience of freedom that belongs to it, the mind acquires knowledge of itself as a product of the Absolute. Jules Lachelier (1832–1918) later gave a vivid and ironical description's of Cousin's method and objectives:

> In the "facts of the will" he thought that he could reach free will, as a permanent power and condition of all consciousness, which is in us the person or the self; in the "rational facts" consciousness seemed to him to raise itself, so to say, above itself to be identified with reason into absolute truth, as it exists both in god and in the universe. Having reached these peaks, nothing would prevent Mr. Cousin from going farther than the boldest kind of metaphysics: for one time he thought that he had demonstrated, through Condillac's method, Schelling's philosophy.[7]

12.2 Comte and Lachelier: Denying the Premise

Auguste Comte, the founder of the positivist school, resisted the inference from psychology to metaphysics by simply denying the spiritualist premise that the mind is knowable by a special method of introspection. In his *Dioptrics*, Descartes (following Aquinas) had denounced the homunculus fallacy: if the mind could observe itself, it would need another observer to observe itself observing and so on.[8] Comte denounces the same fallacy in the Idéologues and in Maine de Biran's attempt to found a science of the mind upon introspection.[9] For him, psychology can only consist of a physiological study of the brain and of a social and historical study of the products of the human mind, i.e., the history of institutions. There are no *facts* of the inner sense, and thus there is no introspective psychology from which metaphysics can derive.

This was also Antoine August Cournot's approach. Like Comte, he pointed out the arbitrariness of the "primitive fact"[10] and denounced the pretensions of the spiritualist philosophers who tried to derive so-called metaphysical truths from psychology. His castigation of his contemporaries' psychologism has much in common

---

[7] Jules Lachelier, "Psychologie et métaphysique," in *Œuvres de Lachelier* (Paris: Alcan, 1896), 105.

[8] AT 6:130. As Cournot reminds us, already Cicero observed: "*Ut oculus, sinc animus, se non videns, alia cernit*," see Antoine August Cournot, *Essai sur les fondements de nos connaissances et sur les caractères de la critique philosophique* (Paris: Vrin-CNRS, 1975 [1851]), 437.

[9] Auguste Comte, *Cours de philosophie positive* (Paris: Rouen frères, 1830; repr. Paris: Hermann, 1975), see especially the 45th lecture.

[10] Cournot, *Essai*, Ch. 23.

with the critique that one finds later in the writings of the German Neo-Kantians and
Husserl at the end of the nineteenth century:

> Apart from empirical psychology which probably belongs to anthropology, [. . .] there is
> another psychology which does not need that apparatus of observations, this slow accumu-
> lation of detailed facts, and which one must no more classify among the sciences of observa-
> tion than arithmetic or geometry, although it relies on some observable and observed facts,
> a condition without which any science would be chimerical. It is clear that one can study
> the conditions of conclusive reasoning, classify our ideas into various categories, expound
> the rules of good method, discuss the value of various proofs and inductions, invoke the
> principles of morals, and further the applications to various species without investigating
> how, under what conditions, in virtue of which forces, and by which natural springs, the
> notions, ideas, rules and principles made their appearance in the mind.[11]

During the second part of the nineteenth century, most members of the spiritual-
ist school renounced the pretensions of Cousin's school to found metaphysics on
psychology. But they did not withdraw from the position that metaphysics, in some
sense, has to be a branch of psychology or psychology a branch of metaphysics.

Henri Bergson, as we shall see, took the first course and argued that metaphysics
is a part of psychology. Lachelier developed the second alternative and argued
that psychology belongs to metaphysics. In his long article "Psychology and Meta-
physics," he tries to arbitrate the disagreement between the partisans of introspec-
tive psychology and those of naturalistic physiology. Against Biran and Cousin, he
contends that psychology is not a form of first philosophy that takes us from facts
about consciousness to metaphysical conclusions. Against the naturalists, he argues
that the mind is not just the object of physiology. Lachelier's main inspiration is
Kantian, and his critique of introspectionism comes quite close to Kant's refutation
of the paralogisms of pure reason. We cannot *know* the self from introspection. But
we can understand it through reflection and the affirmation the self affirms freedom:

> The inner man is double, and there is nothing surprising that he can be the object of two
> sciences [psychology and physiology] which complete each other. Psychology has con-
> sciousness as its proper domain: it knows about thought only the light which it spreads on
> sensation: the true science of thought itself, the light in itself, is metaphysics.[12]

12.3 Bergson: From Intuition to Being

At the time when Lachelier was writing his treatises, psychology was born as a
science in France. Its main advocate, Theodule Ribot (1839–1916), the editor of the
Revue philosophique, was an experimentalist and a naturalist who tried to under-
stand psychological phenomena through the study of neurophysiology. In his lucid
presentation of French philosophy in *Mind*, Ribot derided the French spritualists'
attempt to reach the Absolute through reflection and defended a more modest and

---

[11] *Ibid.*, 440.

[12] Lachelier, "Psychologie et métaphysique," *Œuvres de Lachelier*, 172–73.

more scientific psychology, which he also practiced.[13] The very year when William James published his *Principles of Psychology* (1889), Henri Bergson published his first book, *Essai sur les données immédiates de la conscience*.[14] An interesting feature of Bergson's work, which distinguishes it from the approaches of the other French spiritualists, is the emphasis that it gives to scientific data of psychology and neurology. His book is full of references to physiological optics, to Gustav Fechner's psychophysics, to studies of aphasia and brain lesions, and to other experimental works. In this way, he raises philosophical psychology to a new level which is absent from the writings of his spiritualist predecessors. His main concern, however, is to reinterpret the data of the scientific psychology of his time in order to defend his own philosophical views, which are distinctively dualistic and spiritualistic. Thus, in the *Essay* he attempts to show that mental life cannot be subjected to quantitative measurement and analysis, and that psychological time is misunderstood when it is modelled on a spatial extension. In *Matter and Memory* (*Matière et mémoire*, 1896), Bergson criticises Ribot's view, according to which memories are located in the nervous system and not in the soul.[15] Although he pretends to overthrow the traditional divisions between matter and mind, self and body, or idealism and realism, his own distinction between two kinds of memory – memory as habit, which is mainly bodily and adaptive, and "pure" memory, which is coextensive with consciousness and irreducible to brain processes – is highly dualistic in spirit. According to Bergson, the brain is merely a sensori-motor organ devoted to the regulation of action and not a processor of information or knowledge.

When Bergson comes to describe the method of metaphysics, he characterises it as the method of intuition. Intuition as the proper tool for philosophy is opposed to analysis, the proper tool for science. By "intuition" Bergson means something which is both broader in scope and richer in effects than a simple introspective certainty. The work of intuition, according to Bergson, consists in trying to get into the very nature of the phenomena, as they are given in their flux and in their becoming – what he calls *durée,* durance.[16] In contrast, scientific psychology freezes or solidifies the passing of time and the perpetual movement of the psychic life. When describing the work of intuition, Bergson comes up with the following account:

> Suppose that we place ourselves, by an effort of intuition, within the concrete flow of durance. No doubt there will be no logical reason to posit multiple and varied durances. Conceivably there could be no other durance but our own, just as there might be no other colour in the world but orange, for instance. But just as colour based consciousness, which has an internal affinity with orange instead of perceiving it externally, would feel itself caught between red and yellow, we would even, perhaps, sense unconsciously beneath this colour a whole spectrum displaying the continuum between red and yellow, so the intuition of our durance, so far from leaving us suspended in the void as pure analysis would do,

[13] See Theodule Ribot, "Philosophy in France," *Mind* 2 (1877), 366–86.

[14] See Henri Bergson, *Essai sur les données immédiates de la conscience* (Paris: Alcan, 1889). English translation Henri Bergson, *Time and Free Will: An Essay on the Immediate Data of Consciousness*, trans. F. L. Pogson (Montana: Kessinger Publishing Company, 1910).

[15] Henri Bergson, *Matter and Memory*, trans. N.M. Paul and W.S. Palmer (New York: Zone Books, 1990). Originally published as *Matière et mémoire*, Paris, 1896.

[16] I have adopted F.C.T. Moore's translation of "durée" by "durance".

puts us into contact with a whole continuum of durances which we have to try to follow at a lower or higher level; in both cases we can expand ourselves indefinitely by a more and more violent effort, in both cases we transcend ourselves. In the first case, we proceed in a more fragmented durance, whose palpitations, more rapid than our own, and dividing up simple sensation, dilute its quality into quantity. The limit of this would be the purely homogenous, the pure *repetition* which we define materially. Going in the other direction, we go towards a durance which is more and more tense, contracted and intensified; the limit here would be eternity.[17]

This is meant to be a part of what Bergson calls "true empiricism", which attempts to "get closer and closer to the original itself, to dive deeper into it, and through a sort of *spiritual auscultation*, to feel the palpitation of its soul."[18]

Bergson's notion of intuition differs both from the traditional rationalist concept of intuition and from the introspective understanding that one finds in Maine de Biran and his followers in the French tradition. It differs from the rationalist concept because it is not *intellectual*. He makes this clear when he criticises Kant for having missed the true nature of intuition:

> It is not necessary, to reach intuition, to transport oneself outside the domain of the senses and consciousness. Kant's mistake was to believe that it is. After having proved through decisive arguments that no dialectical effort will ever introduce us into the hereafter and that an effective metaphysics will necessarily be an intuitive metaphysics, he added that we necessarily miss this intuition and that such a metaphysics is impossible. This would be so, indeed, if there were neither any other time nor any kind of change distinct from those that Kant had considered and with which we deal anyway, because our typical perception could neither get out of time nor grasp anything other than change. But the time in which we are necessarily placed and the change to which we are typically subjected, are a time and a change that our consciousness has reduced to ashes in order to facilitate our action upon things. Let us undo what they did, let us bring back our perception to its origins, and we shall have a knowledge of a new kind without having recourse to new faculties.[19]

Intuition should both start from ordinary perception and bring us to the heart of being, which is both duration and *life*. Unlike the concept of intuition that is used by introspective psychology from Maine de Biran to the associationists, this is not any kind of *analysis* or *decomposition* of thought, which would necessarily be abstract. On the contrary, intuition is claimed to grasp the concrete flux of duration in itself, without analysing it into any parts. Intuition-based psychology, as we could call it, is supposed to be true metaphysics, in opposition to the kind of psychology which is practiced and recommended by analytic psychologists such as John Stuart Mill or Hyppolite Taine, "who could not content themselves with being simply psychologists in psychology" and who, although they were psychologists through their method, were metaphysicians by the object which they set to themselves.[20] Analysis produces bad metaphysics, intuition provides good psychological metaphysics.

---

[17] Henri Bergson, *Introduction à la métaphysique* (1903), in *Œuvres*, Edition du Centenaire (Paris: PUF, 1991), 1418–19, trans. Francis Charles Timothy Moore in his *Bergson: Thinking Backwards* (Cambridge: Cambridge University Press, 1996), 92–3.

[18] Bergson, *Œuvres*, 1408.

[19] Henri Bergson, "L'intuition philosophique," in *Œuvres*, 1364.

[20] Bergson, *Œuvres*, 1406.

Psychological metaphysics, based on intuition, has not only the positive result of bringing us to the heart of life and being, but it also has a negative or critical correlate in the denouncement of the philosophical illusions which arise when the metaphysical problems are ill-posed. The most well known example of the critical aspect of Bergson's analyses is his famous attack upon the metaphysical notion of possibility. He argues that we represent possibilities to ourselves as states of the world that are not actual but that could be realised. Possibility is nothing but a mirage or an illusion created by projecting into the past an image of the present states of the world and supposing that it has been actualised. There is nothing possible, everything is actual, in the present of consciousness. We merely *project* into the past – and into the future – our present perception in a "retrograde motion".[21] This illusion of retroaction of the past, which creates the ghost of the possible, is also responsible for our conception of time as a structure of branching possibilities and for our conception of free will as a choice between possibilities.

This negative strand of Bergson's thought has often been compared to the kind of criticism of metaphysics that has been presented by analytic philosophers, for instance, when pointing out category mistakes.[22] But there is a crucial difference between these two critical approaches. For Bergson, the factitiousness of philosophical problems and the falsity of a number of metaphysical theses do not stem from a failure to respect grammatical or logical distinctions but from a failure to place oneself in the position required by the method of intuition: inside the true movement of duration and time. Given that this method is a sophisticated form of psychological introspection, as we have argued, it follows that the intuition of duration is the true criterion for the meaningfulness of any metaphysical concept or problem. Metaphysical concepts, such as 'possibility', 'order', or even 'free will', are misconceived when they are approached from a point of view that is alien to the perspective of intuition and when they are mere products of psychological projection.

It is at this point that the principal difficulty of the spiritualist strategy arises. For, on the one hand, its project is to reach metaphysical *facts* from a psychological standpoint – the standpoint of introspection and reflection or "intuition"; but, on the other hand, it makes these facts relative to the psychological inquiry. How can one reach metaphysical *objectivity* or metaphysical facts from a point of view which is *subjective* in essence? This is the first part of the difficulty. The other follows immediately. If the so-called objective metaphysical facts are actually constitutively psychological in the sense of being produced by some process of introspection, reflexion, or intuition, then they are not objective at all. They are actually

---

[21] Henri Bergson, "La pensée et le mouvant," in *Œuvres*, 1331–345. I have analysed these points in Pascal Engel, "Plenitude and Contingency: Modal Concepts in Nineteenth Century French Philosophy," in Simo Knuuttila (ed.), *Modern Modalities* (Dordrech: Kluwer, 1988) and Pascal Engel, "Lagneau entre Ribot et Bergson," in *Jules Lagneau: cours integral 1886–7* (Dijon: CRDP de Bourgogne, 1997), 90–100.

[22] Such a reading of Bergson is proposed by Francis Charles Timothy Moore, *The Psychology of Maine de Biran* (Cambridge: Cambridge University Press, 1970).

psychological facts. This explains why Bergson can attempt what looks very much like a psychological *debunking* of the notion of possibility. Such debunking was not uncommon in his time, and it has some affinities with Nietzsche's attempt to criticise metaphysics through his own version of the psychology of sentiments: "We psychologists of the future . . ." In other words, French spiritualism attempted to preserve metaphysics as a science of being, but it ended up dissolving it in the psychological washbasin, to use Frege's terms when he criticised Husserl's early psychologism.

12.4 Conclusion

When the French spiritualists conceived their project of discovering metaphysical truths by an introspective inquiry into our minds, they did not aim at what the empiricist philosophers in the eighteenth century called the origin of our ideas about space, time, object, possibility, or the self. Neither did they aim at any form of psychologisation of metaphysics or at any anthropologisation of concepts, as the one attempted by the German "psychologists" philosophers of the same time, such as Beneke, Fries, Herbart, and others, or by heirs of British empiricism, such as John Stuart Mill or Alexander Bain.[23] They did not work to bury metaphysics, but to praise it. They intended to preserve and value the metaphysical truths about being, the self, and substance, which the rationalist philosophers had attempted to discover through intellectual intuition. But at the same time, they did not accept the assumptions of rationalistic thinkers, or their apriorism. They hoped to get to the highest regions of Being and Spirit, but starting from a psychological inquiry that would take into account both the operations of the senses and the power of conscious reflexion. They accepted Kant's critique of metaphysics and his rejection of empiricism, but they also dreamt of a "true empiricism" which would allow the mind to communicate directly with the nature of the self and with the very essence of life. They also rejected psychology as an objective, third-personal, and scientific inquiry into the mind.

The result of the inquiries of the spiritualists was bound to look like a hybrid mixture of psychology and metaphysics, congenial to what Cousin called "eclectism" and to what Ravaisson called "spiritualistic positivism" but in fact much closer to a form of psychologism than they claimed it to be. It is indeed, in all respects, a form of idealism: what is experienced is supposed to work as a guide to what exists, what is conceivable is taken as a guide to what is possible, and true reality is spiritual in nature. Bergson's critique of the illusion of possibility, for example, although it could be interpreted as a form of what is today known as actualism in modal metaphysics, is much closer to a psychological account of the genesis of our concept of possibility. Bergson's main argument is that we have a mirage *representation* of

---

[23] On the history of psychology in the nineteenth century see Riccardo Martinelli, *Misurare l'anima, filosofia e psicofisica da Kant a Carnap* (Macerata: Quodlibet, 1999).

the present in the past, which produces in us the *mental* illusion of possibles that would be actualised, and that we can only *conceive* of a world which is full of being.

This is not to say that the project of attempting to derive essential properties of things from introspective data is in itself absurd or contradictory. At the time when Bergson was writing, Brentano formulated his own project of a descriptive or phenomenological psychology that was supposed to proceed to the essence of mental states, starting from an introspective investigation into our mental acts. What differentiates French spiritualism from an enterprise of this sort? Firstly, the answer lies in the fact that Brentano's psychology was meant to be *analytic* in precisely the sense that Bergson rejected. Secondly, it was embedded in a metaphysics which was *independent of*, and not *derived* from, the introspective investigation that constitutes descriptive phenomenology. Thus, according to Brentano psychology depends on metaphysics, whereas according to French Spiritualism metaphysics depends on psychology.

Similarly, we should not criticise the French spiritualists for having emphasised the faculty of intuition but for having relied too heavily and too uncritically on it. Bertrand Russell famously criticised Bergson's notion of intuition and accused it of mysticism.[24] But it does not follow that there is no place for an epistemological theory of intuition and its role in knowledge. On the contrary, a number of contemporary philosophical enterprises are based upon intuition.[25]

Neither should my critical remarks about French spiritualism be understood as restating the objection that Father Bourdin presented to Descartes, claiming that conceivability does not imply possibility. It is questionable wheter we can derive any propositions about what is possible from what we sense or feel, or from what we imagine, but it is more probable that we can draw conclusions about the metaphysical possibility from what can *conceptually* be conceived and thought.[26] Some serious contemporary projects in metaphysics are based on the idea that conceptual analysis *can*, under certain constraints, ground conclusions about metaphysics. This project can also be understood as a project of analysing the basic concepts of folk-psychology – understood as a functionalist theory of mind – and deriving certain metaphysical conclusions from this basis.[27] I do not of course mean to endorse such a project but merely want to point out that there is no doubt that it makes sense. It is equally obvious that such project has very little to do with the kind of developments described in this chapter.

---

[24] See Bertrand Russell, *Mysticism and Logic* (London: Allen and Unwin, 1917).

[25] Most prominently Roderick M. Chisholm's foundationalism, see his *The Foundations of Knowing* (Minneapolis: University of Minnesota Press, 1981) and more recently Michael R. De Paul and William Ramsey, *Rethinking Intuition* (Lanahm: Rowman and Littlefield, 1998).

[26] On this contemporary debate, see the essays in Gendler and Hawthorne (eds.), *Conceivability and Possibility*.

[27] See Frank Jackson, *From Metaphysics to Ethics* (Oxford: Oxford University Press, 1998).

Chapter 13
Philosophy, Psychology, Phenomenology

Dan Zahavi

Abstract The chapter investigates the relation between philosophy, psychology, and phenomenology. First, it gives a brief account of Husserl's criticism of psychologism in his *Logical Investigations* (*Logische Untersuchungen*). Then it looks at Husserl's later distinction between a phenomenological psychology and a transcendental phenomenology. And finally, the chapter discusses the difference between a (neo-)Kantian and a phenomenological concept of transcendental philosophy.

Keywords Phenomenology · phenomenological psychology · transcendental philosophy · Husserl · Kant

Husserl starts his lecture course *Phenomenological Psychology* (*Phänomenologische Psychologie*) from 1925 with the following historical overview. During the nineteenth century, psychology underwent a tremendous development. Through the collaborative effort of scientists like Müller, Weber, Volkmann, Helmholtz, Hering, Fechner and Wundt, psychology was established as a truly scientific enterprise. But although experimental psychology quickly gained results, it was a psychology that looked to the natural sciences, especially to physiology, for guidance, it was a psychology that – in Husserl words – could not withstand the temptation of naturalism, and it was soon met with harsh criticism.[1] Could psychology of this kind really account for the experiential dimension, did it at all take subjectivity into account? In his text *Ideen über eine beschreibende und zergliedernde Psychologie* (1894), Wilhelm Dilthey argued that psychology had become much too influenced by the natural sciences and that it should to a larger extent orient itself towards the humanities. Dilthey wished to preserve the difference between the explanation offered by the natural sciences and the understanding obtained by the humanities, and he criticized the naive attempts to naturalize the latter. Dilthey distinguished an explanatory

D. Zahavi
Center for Subjectivity Research, Faculty of Humanities, University of Copenhagen

[1] Edmund Husserl, *Phänomenologische Psychologie*, Husserliana IX (Den Haag: Martinus Nijhoff, 1962), 3–5.

or constructive psychology from his own descriptive or analyzing psychology. The former was committed to the ideals of natural science. It was atomistic and based on hypotheses and inferences. But by proceeding in such a manner, it was also incapable of grasping the life of consciousness. It failed to realize that the unity of psychic life is an experienced unity and not simply an inferred and postulated one.[2] However, according to Husserl, Dilthey's counter-attack was not completely successful. On the one hand, Dilthey's own alternative remained too wedded to an introspective methodology. It only dealt with particularities, and was unable to offer universal insights and formulate proper psychological laws.[3] On the other hand, its criticism of experimental psychology was not sufficiently worked out, and it was countered in 1895 by a brilliant and devastating article by Hermann Ebbinghaus, who had pioneered in the development of experimental methods for the measurement of rote learning and memory. At the turn of the century, the battle consequently seemed to have been won by the naturalists. But then, of course, *Logical Investigations* (*Logische Untersuchungen*, 1900–1901) was published, and as Husserl remarks, this changed everything.[4]

In the following, I intend to investigate the relation between philosophy, psychology, and phenomenology. My presentation will have the following parts: First, I will give a brief account of Husserl's criticism of psychologism in *Logical Investigations*. Then I will look at Husserl's later distinction between a phenomenological psychology and a transcendental phenomenology. And finally, I will discuss the difference between a (neo-)Kantian and a phenomenological concept of transcendental philosophy.

13.1 Husserl's Anti-Psychologism

In the preface to the first edition of *Logical Investigations*, Husserl writes that after the publication of his first book *Philosophy of Arithmetic* (*Philosophie der Arithmetik*, 1891) he came to doubt whether it was possible to reconcile the objectivity of science with a psychologistic foundation of logic.[5] More specifically, the task Husserl set himself in the first part of *Logical Investigations*, the part entitled *Prolegomena to Pure Logic* (*Prolegomena zur reinen Logik*), was to distinguish pure logic from logic as a normative science, as well as to reject the claim that logic is rooted in the empirical nature of consciousness.

Husserl's main target was the position known as *psychologism*. In its view, epistemology is concerned with the cognitive nature of perceiving, believing, judging, and knowing. All of these phenomena, however, are psychical phenomena, and it therefore seems rather obvious that it must be the task of psychology to investigate and

---

[2] Husserl, *Phänomenologische Psychologie,* 7–8, 14.

[3] *Ibid.,* 16.

[4] *Ibid.,* 20.

[5] Edmund Husserl, *Logische Untersuchungen. Erster Band. Prolegomena zur reinen Logik,* Husserliana XVIII (Den Haag: Martinus Nijhoff, 1975), 6–7.

explore their structure. This also holds true for our scientific and logical reasoning, and ultimately logic must therefore be regarded as a part of psychology, and the laws of logic as psychological regularities whose nature and validity must be empirically investigated.[6] Thus, psychology turns out to provide logic with its theoretical foundation.[7]

According to Husserl, however, psychologism commits the error of ignoring the fundamental difference that exists between the domain of *logic* and *psychology*. Logic (as well as for instance mathematics and formal ontology) is not an empirical science, and it is not at all concerned with factually existing objects. On the contrary, it investigates ideal structures and laws, and its investigations are characterized by certainty and exactitude. In contrast, psychology is an empirical science that investigates the factual nature of consciousness, and its results are therefore characterized by the same degrees of vagueness and probability that marks the results of all the other empirical sciences.[8] To reduce logic to psychology is consequently a regular category mistake that completely ignores the ideality and apodicticity characterizing the laws of logic.[9] These features can never be founded in or explained by reference to the factual-empirical nature of the psyche.

Whereas a psychical process elapses in time and has a beginning and an end, this does not hold true for the logical principles or mathematical truths which are known.[10] When one speaks of a law of logic or refers to mathematical truths, to theories, principles, sentences, and proofs, one does not refer to a subjective experience with a temporal duration, but to something atemporal, objective, and eternally valid. Although the principles of logic are grasped and known by consciousness, we remain conscious of something *ideal* which is irreducible to the conscious process of knowing.

Had ideality really been reducible to and therefore susceptible to the influence of the temporal, real, and subjective nature of the psychical process, it would have been impossible to repeat or share meaning, just as it is impossible to repeat a concrete psychical process the moment it has occurred, not to speak of sharing it with others. (We can of course effectuate a *similar* process, but similarity is not identity). But had this really been the case, scientific knowledge as well as ordinary

---

[6] *Ibid.,* 64, 89.

[7] Psychologism – a position often believed to have been long dead and buried – has recently had a revival, this time in the form of what might be called *neurologism*. As the well-known neuroscientist Semir Zeki wrote in a recent article: "My approach is dictated by a truth that I believe to be axiomatic – that all human activity is dictated by the organization and laws of the brain; that, therefore, there can be no real theory of art and aesthetics unless neurobiologically based" (Semir Zeki, "Neural Concept Formation & Art: Dante, Michelangelo, Wagner", *Journal of Consciousness Studies* 9:3 (2002), 54). I think the problematic nature of this proposal becomes obvious the moment one simply replaces art and aesthetics with other human activities, such as astrophysics or archaeology.

[8] Husserl, *Logische Untersuchungen*, 181.

[9] *Ibid.,* 79–80.

[10] Edmund Husserl, *Einleitung in die Logik und Erkenntnistheorie*, Husserliana XXIV (Den Haag: Martinus Nijhoff, 1984), 141.

communication and understanding would have been impossible.[11] Thus, Husserl can accuse psychologism of entailing a self-refuting skepticism. To attempt to *naturalize* ideality by reducing it to psychical processes is, according to Husserl, to undermine the very possibility of any theory, including psychologism itself.

The central and positive task of the *Prolegomena* was to show that objectivity and scientific knowledge presuppose ideality. But even if it is impossible to reconcile scientific objectivity with a psychologistic take on logic, one is still confronted with the apparent paradox that objective truths are known in subjective experiences of knowing. And as Husserl points out, this relation between objectivity and subjectivity has to be investigated and clarified if we wish to attain a more substantial understanding of the possibility of knowledge.[12] The answers to the fundamental questions that we find in epistemology and in the theory of science consequently call for an "unnatural" change of interest. Instead of merely paying attention to the objects, Husserl urges us to reflect on, thematize, and analyze the structure of experience and intentionality, for it is only in this manner that we will be able to reach an understanding of the relation between the act of knowing and the object of knowledge.[13]

Despite Husserl's strong criticism of psychologism, his interest in the fundamental problems of epistemology consequently made him return to consciousness. But this did not entail a return to psychologism.[14] First of all, there is no attempt to *reduce* the object to the experience, but only an attempt to understand the object in *relation* or *correlation* to the experience. Secondly, Husserl wants to *understand* and *describe* the a priori structure of these experiences, he is not interested in a naturalistic *explanation* that seeks to uncover their biological genesis, neurological basis, psychological motivation, or the like. Although Husserl himself had in the first edition of *Logical Investigations* been imprudent enough to characterize phenomenology as a *descriptive psychology*, he soon realized that this was a serious mistake.[15] Husserl was neither interested in an analysis of the psycho-physical constitution of man, nor in an investigation of empirical consciousness, but in an understanding of that which in principle and intrinsically characterizes perceptions, judgments, feelings etc.[16]

---

[11] Husserl, *Logische Untersuchungen*, 194.

[12] Husserl's distinction between the ideal and the real are in many ways similar to Frege's distinction. But the very important difference between the Husserlian and the Fregean criticism of psychologism is that Husserl believed it to be necessary to continue this criticism by way of an analysis of intentionality, and this interest in the structures of subjectivity and the first-person perspective was not shared by Frege.

[13] Husserl, *Einleitung in die Logik und Erkenntnistheorie*, 14.

[14] Husserl, *Logische Untersuchungen*, 11; Edmund Husserl, *Logische Untersuchungen II*, Husserliana XIX/1–2 (Den Haag: Martinus Nijhoff, 1984), 535.

[15] Edmund Husserl, *Aufsätze und Rezensionen (1890–1910),* Husserliana XXII (Den Haag: Martinus Nijhoff, 1979), 206–08.

[16] Husserl, *Logische Untersuchungen II*, 23, 357; Husserl, *Aufsätze und Rezensionen*, 206–08; Husserl, *Phänomenologische Psychologie,* 39.

It could be argued, however, that even if Husserl's phenomenological analysis of intentionality might not constitute a relapse into psychologism, it should certainly be appreciated as a new type of psychology. In fact, the best way to interpret the notorious phenomenological method is precisely to interpret it as a new psychological method. Or is it? In a manuscript entitled *Phänomenologie und Psychologie* from 1917, Husserl raises the following question: Why introduce a new science entitled 'phenomenology' when there is already a well-established explanatory science dealing with the psychic life of humans and animals, namely psychology. More specifically, psychology is a science of naturalized consciousness. And could it not be argued that a mere description of experience – which is supposedly all that phenomenology can offer – does not constitute a viable scientific alternative to psychology, but merely a – perhaps indispensable – descriptive preliminary to a truly scientific study of the mind?[17] As Husserl remarks, this line of thought has been so convincing that the term "phenomenological" is being used in all kinds of philosophical and psychological writings to describe a direct description of consciousness based on introspection.[18] The parallel to the contemporary discourse is quite striking. Currently, the term "phenomenology" is also increasingly being used by analytical philosophers and cognitive scientists to designate a first-person description of what the "what it is like" of experience is really like. And against that background, it might be difficult to understand why phenomenology should not simply be classified as a kind of psychology, or even as a form of introspectionism.

It is undeniable that phenomenology has affinities with psychology insofar as both disciplines are interested in consciousness. But as Husserl also points out, although the distinction between a phenomenological and a psychological investigation of consciousness can be difficult to draw, and might at first even appear as an unnecessarily subtle distinction, we are in the end confronted with an absolute crucial nuance that is fundamental to the very possibility of doing philosophy.[19]

Husserl takes psychology to be an empirical science about the *nature* of the psyche, and therefore to be a science about psychical life understood as a real occurrent entity in the natural world.[20] By contrast, phenomenology is not empirical, but eidetic and a priori. And even more importantly, phenomenology is not interested in consciousness as a natural occurrence. Phenomenology seeks to describe the experiential structures in their phenomenal purity and does not, as Husserl puts it, psychologize them, that is, it does not objectify and naturalize them.[21] But although phenomenology and psychology differ, this does not make them unrelated. Husserl characterizes phenomenology (a phenomenology not misled by naturalistic prejudices, as he adds) as a foundation and presupposition for a truly scientific

---

[17] Edmund Husserl, *Aufsätze und Vorträge (1911–1921),* Husserliana XXV (Den Haag: Martinus Nijhoff, 1987), 102.

[18] *Ibid.,* 103.

[19] Husserl, *Einleitung in die Logik und Erkenntnistheorie*, 211.

[20] Husserl, *Aufsätze und Vorträge (1911–1921)*, 75.

[21] *Ibid.,* 117.

psychology.[22] Experimental psychology has produced striking results, but it pre-supposes something that it does not itself deliver, namely a proper investigation of (subjective) consciousness.[23] In fact, as Husserl remarks, the new scientific psychol-ogy is basically a psychology that has lost sight of consciousness.[24] This is where phenomenology can help out.

Given what has been said so far, it seems natural to conclude as follows: There is more to phenomenology than simply a compilation of introspective reports. What needs to be realized is that phenomenology can contribute with refined and so-phisticated analyses of the *a priori* structures of consciousness. This is where phe-nomenology differs from mere empirical psychology.

But this conclusion represents yet another misunderstanding. As Husserl writes, "The eidetic analyses of experiential consciousness, the consciousness of outer ex-perience and each experience, and thus continued the eidetic analyses of all kinds of consciousness may be carried far: we still remain on the ground of psychology".[25] In other words, if all phenomenology could do was to contribute with refined *a pri-ori* analyses of consciousness, phenomenology would not differ from a (rigorously implemented) psychology, and it would consequently be superfluous.

So we are back where we started. We are still faced with the following ques-tion: Is there a difference between a phenomenological and a (folk-)psychological investigation of consciousness, and if so, where is this difference to be located?

13.2 Phenomenological Psychology and Transcendental Phenomenology

Logical Investigations heralded the birth of a new method for studying conscious-ness, a method called *phenomenology*.[26] The aim was to explore the intentional structures involved in our perception, thinking, judging, etc. This might seem like a simple continuation of the project commenced by Brentano in his *Psychology from an Empirical Standpoint* (*Psychologie vom empirischen Standpunkt*, 1874). But al-though Brentano should be praised for his rediscovery of the concept of intentional-ity, his analysis of intentionality remained – as Husserl points out – naturalistic and psychological, whereas Husserl's own analysis was neither.[27] Thus, it is important to realize that the stated purpose of *Logical Investigations* was not to establish a new foundation for psychology, but rather to provide a new foundation for episte-mology. And although this was not yet fully realized in *Logical Investigations*, such

---

[22] Husserl, *Einleitung in die Logik und Erkenntnistheorie*, 383–84; Husserl, *Aufsätze und Vorträge (1911–1921)*, 39.

[23] Husserl, *Aufsätze und Vorträge (1911–1921)*, 19.

[24] *Ibid.,* 104.

[25] *Ibid.,* 104.

[26] Husserl, *Phänomenologische Psychologie,* 28, 302.

[27] *Ibid.,* 37, 310.

a task would ultimately necessitate a clear distinction between two quite different perspectives on consciousness, a psychological and a transcendental one.[28]

In the long essay *Philosophy as a Rigorous Science* (*Philosophie als strenge Wissenschaft*) from 1911, Husserl explicitly contrasts his own phenomenological investigation of consciousness with a natural scientific account of consciousness, and he argues that the attempt to naturalize consciousness has not only failed, but that it is fundamentally flawed.[29] That Husserl felt strongly about this issue is clearly expressed in the following letter to the neo-Kantian Heinrich Rickert from 1915:

> During the last decade I feel closely connected to the leaders of the schools of German Idealism; together we both fight a common enemy, the naturalism of our time. We both serve the same gods, each in his own way, and since all of us consider this service a solemn and holy matter to which we have devoted our entire life, it will carry its own necessity and will be indispensable for the progress of philosophy.[30]

Husserl's criticism of naturalism can provide us with a clue to the real difference between phenomenology and psychology, for one of the main problems with naturalism is, according to Husserl, that it hinders a recognition of transcendental subjectivity. When I perform a psychological reflection, I am interpreting the act reflected upon as a psychical process, that is, as a process occurring in a psycho-physical entity that exists in the world. This type of self-experience is just as worldly an experience as, say, the experience of a chair or a table, and if one asks whether it can provide us with an adequate understanding of subjectivity the answer is no. Psychological reflection or introspection presents us with a constituted, objectified, and naturalized subject. It does not provide us with an access to the constituting, transcendental, dimension of subjectivity.[31]

It would, of course, be something of a slight exaggeration to claim that the notion of transcendental subjectivity is universally accepted in contemporary philosophy, but I think that much of the criticism is based on something that approaches a complete misunderstanding of the term. Often transcendental subjectivity is taken to be some kind of other-worldly, ghostly, homunculus. Confronted with such ignorance, it is crucial to demythologize the notion.

The empirical subject and the transcendental subject are not two different subjects, but rather two different perspectives on one and the same subject.[32] It is

---

[28] *Ibid.,* 42.

[29] Husserl, *Aufsätze und Vorträge (1911–1921)*, 17, 41.

[30] Edmund Husserl, *Briefwechsel,* Husserliana Dokumente III/1–10 (Dordrecht: Kluwer Academic Publishers, 1994), 5/178.

[31] Edmund Husserl, *Formale und Transzendentale Logik*, Husserliana XVII (Den Haag: Martinus Nijhoff, 1974), 290; Edmund Husserl, *Erste Philosophie (1923/24). Zweiter Teil. Theorie der phänomenologischen Reduktion,* Husserliana VIII (Den Haag: Martinus Nijhoff, 1959), 71; Edmund Husserl, *Erste Philosophie (1923/24). Erster Teil. Kritische Ideengeschichte,* Husserliana VII (Den Haag: Martinus Nijhoff, 1956), 269; Edmund Husserl, *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie. Eine Einleitung in die phänomenologische Philosophie,* Husserliana VI (Den Haag: Martinus Nijhoff, 1954), 255, 264.

[32] Husserl, *Phänomenologische Psychologie*, 294.

a difference between conceiving of the subject as an object in the world, and conceiving of the subject as a subject for the world, i.e., as a meaning-bestowing and world-disclosing subject of intentionality. The whole thrust of the phenomenological analysis is to unearth the latter, thereby disclosing what phenomenology takes to be a *non-psychological* dimension of consciousness.[33]

Husserl's rejection of any straightforward identification of phenomenology with psychology was fully shared by Heidegger. Not only does Heidegger categorically reject that his own analysis of the existential structures of Dasein is a psychological analysis, but he also writes that the attempt to interpret Husserl's investigations as a kind of descriptive psychology completely fails to do justice to their transcendental character.[34] In fact, as Heidegger adds, phenomenology will remain a book sealed with seven or more seals to any such psychological approach.[35] For both Husserl and Heidegger phenomenology differs from psychology by not simply accepting the ontological (or metaphysical) presuppositions of the natural attitude. We should realize, as Husserl remarks,

> [. . .] that every historically available scientific discipline and even part of what normally belongs to philosophy, including formal logic, psychology and ethics, conduct their research in a natural – though in a certain way necessary – naivety. All their questions refer to a world which is given to us – with an obviousness belonging to life – prior to all science, but they fail to notice that this pre-givenness conceals a true infinity of enigmatic problems, which are not even noticed from within the natural perspective. I am referring to the transcendental problems, and it is only the truly scientific philosophy which has made their disclosure possible.[36]

Naturalism has denied the existence of a particular philosophical method, and has claimed that philosophy should employ the same method that all strict sciences are using, the natural scientific method. But for Husserl this line of reasoning merely

---

[33] The phenomenologists themselves have vehemently denied that they are engaged in some kind of introspective psychology. For some representative statements, cf. Aron Gurwitsch, *Studies in Phenomenology and Psychology* (Evanston: Northwestern University Press, 1966), 89–106; Husserl, *Einleitung in die Logik und Erkenntnistheorie*, 201–216; Martin Heidegger, *Grundprobleme der Phänomenologie (1919/1920),* Gesamtausgabe Band 58 (Frankfurt am Main: Vittorio Klostermann, 1993), 11–7. One simple argument is that phenomenology must be appreciated as a kind of transcendental philosophy; it is not a kind of empirical psychology. It is not interested in psychical processes but in meaning structures. Another related argument is that introspection is typically understood as a mental operation that enables us to report and describe our own mental states. But strictly speaking phenomenology is not concerned with or based on operations of this kind. Rather, phenomenology is concerned with the phenomena, the appearances, their essential structures, and their conditions of possibility, and phenomenologists would typically argue that it is a metaphysical fallacy to locate these appearances within the mind, and to suggest that the way to access and describe them is by looking inside (*introspicio*) the mind. The entire facile divide between inside and outside is phenomenologically suspect, but this divide is precisely something that the term "introspection" acknowledges and accepts, cf. Dan Zahavi, *Husserl's Phenomenology* (Stanford: Stanford University Press, 2003).

[34] Martin Heidegger, *Sein und Zeit* (Tübingen: Max Niemeyer, 1986), 45–50.

[35] Heidegger, *Grundprobleme der Phänomenologie (1919/1920)*, 15–6.

[36] Edmund Husserl, *Natur und Geist. Vorlesungen Sommersemester 1927*, Husserliana XXXII (Dordrecht: Kluwer Academic Publishers, 2001), 7.

displays that one has failed to understand what philosophy is all about. Philosophy has its own aims and methodological requirements; requirements that for Husserl are epitomized in his notion of *phenomenological reduction*.[37] For Husserl, the reduction is meant to make us maintain the radical difference between philosophical reflection and all other modes of thought. As he wrote already in 1907: "*Thus, the 'phenomenological reduction' is simply the requirement always to abide by the sense of the proper investigation, and not to confuse epistemology with a natural scientific (objectivistic) investigation.*"[38] Every positive science rests upon a field of givenness or evidence that is presupposed but not investigated by the sciences themselves. In order to make this dimension accessible, a new type of inquiry is called for, a type of inquiry that "*precedes all natural knowledge and science and points in a quite different direction than natural science.*"[39] The task of phenomenology is consequently not to describe objects as precisely and meticulously as possible, nor should it concern itself with an investigation of the phenomena in all their ontic diversity. Rather, its true task is to examine the very dimension of appearance or givenness, and to disclose its inner structure and condition of possibility. To thematize the objects in terms of their givenness, validity, and intelligibility calls for a reflective stance quite unlike the one needed in the positive sciences. This, of course, is one reason why the phenomenological attitude has frequently been described as an unnatural direction of thought.[40] But to describe phenomenology as unnatural is of course also to deny any straightforward continuity between phenomenology and positive sciences like psychology.

The positive sciences are so absorbed in their investigation of the natural (or social/cultural) world that they do not pause to reflect upon their own presuppositions and conditions of possibility. The positive sciences operate on the basis of a natural (and necessary) naivety. They take it for granted that the world exists. But this assumption must be philosophically investigated. That such an investigation is required should not, however, be taken as an endorsement of skepticism. That the world exists is, as Husserl writes, beyond any doubt. But the great task is to truly understand this indubitability (which sustains life and positive science) and to clarify its legitimacy.[41] That is precisely the aim of the phenomenological investigation of intentional consciousness. As Husserl writes:

> The task that now arises is how to make this correlation between constituting subjectivity and constituted objectivity intelligible, not just to prattle about it in empty generality but to clarify it in terms of all the categorial forms of worldliness, in accordance with the universal structures of the world itself. If we accept the premise that the constitutive functions of

---

[37] Husserl, *Einleitung in die Logik und Erkenntnistheorie*, 238–39.

[38] *Ibid.,* 410.

[39] *Ibid.,* 176.

[40] Cf. Husserl, *Logische Untersuchungen II*, 14.

[41] Edmund Husserl, *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Drittes Buch: Die Phänomenologie und die Fundamente der Wissenschaften,* Husserliana V (Den Haag: Martinus Nijhoff, 1952), 152–53; Husserl, *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie*, 190–91.

consciousness, both active and passive, are actually to be brought to light, functions which make evident to us the meaning and self-verifying being of a world we accept as there, then this task is manifestly a totally different one from that of all positive sciences – and, as compared with all of them, is completely new. For all of these sciences, the intelligible existence of a world is presupposed, and its fundamental knowability, also, to no less a degree. Both of these remain outside the topic.[42]

By adopting the phenomenological attitude we become aware of the givenness of objects. But we do not simply focus on the objects precisely as they are given; we also focus on the subjective side of consciousness, thereby becoming aware of our subjective accomplishments and of the intentionality that is at play in order for the objects to appear as they do. When we investigate appearing objects, we also disclose ourselves as datives of manifestation, as those to whom objects appear.

All phenomenologists have emphasized the importance of the first-person perspective. But why have phenomenologists been so preoccupied with such an examination of subjectivity and selfhood? Had it been a goal in itself, phenomenology would have remained a form of philosophical psychology or anthropology. But the phenomenological interest in subjectivity is not motivated by the relatively trivial insight that we need to include the first-person perspective if we wish to understand mental phenomena. Rather the analysis is transcendental philosophical in nature. The reason why phenomenologists have been so preoccupied with describing and analyzing the fundamental features of subjectivity, be it its structures of intentionality, of embodiment, of temporality, of historicity or of intersubjective embeddedness, is because they have been convinced that a thorough philosophical understanding of the world that we experience and live in must include an investigation of subjectivity. Not in the sense that in order to investigate the world we will first have to investigate subjectivity in order then, indirectly, to reach the world. Rather, the idea is that every phenomenon, every appearance, is always an appearance of something for someone. If we wish to understand how physical objects, mathematical models, chemical processes, social relations and cultural products can appear as they do and with the meaning they have, then we will also have to examine the subject to whom they appear. When we encounter perceived, judged, evaluated objects, a thorough philosophical examination of these objects will lead us to the experiential structures which these modes of appearance are correlated with. We will be led to the acts of presentation, perception, judgment, and valuation, and thereby to the subject (or subjects) which the object as appearing must necessarily be understood in relation to.

Husserl occasionally distinguishes between a phenomenological psychology and a transcendental phenomenology.[43] The task of phenomenological psychology is to investigate intentional consciousness in a non-reductive manner, that is, in a manner that respects its peculiarity and distinctive features. Phenomenological psychology is consequently a form of descriptive, eidetic, and intentional psychology which takes the first-person perspective seriously, but which – in contrast to transcendental

---

[42] Husserl, *Phänomenologische Psychologie,* 336–37.
[43] *Ibid.,* 343.

phenomenology, that is, the true philosophical phenomenology – remains within the natural attitude. In other words, the difference between the two is that whereas phenomenological psychology might be described as a so-called regional-ontological investigation which investigates consciousness for its own sake, transcendental phenomenology is a much more ambitious global enterprise. It is interested in the constitutive dimension of subjectivity, that is, it is interested in an investigation of consciousness insofar as consciousness is taken to be a condition of possibility for meaning, truth, validity, and appearance.

13.3 Two Types of Transcendental Philosophy

Husserl was at first strongly influenced by Brentano's very negative appraisal of Kant. But as he recounts in the following letter to Ernst Cassirer from 1925, subsequent studies made him realize the affinity between his own project and Kant's:

> My own development – which originally was hostile towards Kant but admittedly also insensitive to the proper sense of the Kantian philosophy – linked up with Descartes and the pre-Kantian philosophy of the eighteenth century, and I was of course also influenced by important impulses from Brentano, Lotze and Bolzano. But as I, continually reflecting on the possibility of a presuppositionless ascertainment and absolute accountability, was driven from the foundational problems belonging to the theory of science (which were the questions that were most familiar to me as a mathematician) towards a method of eidetic analysis of consciousness, and as the domain of the origins of all knowledge was disclosed to me through the phenomenological reduction, I had to acknowledge that the emerging science, although with quite different methods, embraced Kant's entire complex of problems (which only now acquired a profound and clear meaning), and that it confirmed and justified Kant's main results in a strictly scientific grounding and limitation.[44]

There is, of course, one place where Kant's influence on Husserl is particularly visible. As Husserl admits in *Erste Philosophie I*, when he decided to designate his own phenomenology as transcendental, he was employing a Kantian concept.[45] But although a reference to Kant is appropriate, and although such a reference can certainly be helpful in the attempt to explain why phenomenology is not simply a form of introspective psychology, one should also be careful not to overlook some rather decisive differences between Kantianism and phenomenology. To put it another way, it would be a mistake to think that transcendental philosophy is all one thing, and to overlook the difference between a Kantian transcendental philosophy and the form of transcendental philosophy we find in phenomenology.

To illustrate this difference, let us take a brief look at Steven Crowell's recent discussion of the work of the neo-Kantian Emil Lask. According to Crowell, Lask emphasized the difference between a positive (empirical or metaphysical) investigation and a transcendental investigation. This difference should be understood as a difference between a straightforward thematization of entities that aims to

---

[44] Husserl, *Briefwechsel*, 5/4.
[45] Husserl, *Erste Philosophie (1923/24). Erster Teil*, 230.

uncover their material properties and a reflective thematization of the truth structure or intelligibility of entities. Lask consequently operates with a duality between the realm of entities (be they physical or metaphysical) on the one hand, and the realm of validity (truth, intelligibility) on the other. This difference between what is and what holds is basically a distinction between the entity and its meaning. This in turn is not an empirical distinction, but rather a transcendental difference between the way one and the same entity can be taken, first in straightforward experience and then again in reflective inquiry. Thus, Lask was quite explicit about the reflective character of transcendental philosophy. Its first principle is not an entity, be it a subject or a substance, but something more fundamental, namely meaning.[46]

Lask's view is to a large extent shared by both Heidegger and Husserl, but for both of the latter (and this is one important area where phenomenology departs from neo-Kantianism) the investigation of meaning has to be grounded in a theory of intentionality, since only such a theory can furnish us with a clarification of evidence and givenness, and this is needed if dogmatism is to be avoided. Thus, in contrast to Lask, Husserl and Heidegger do not believe that a transcendental theory of meaning can be carried through without a simultaneous inquiry into the subjective dimension of meaning disclosure.[47] Or to use a more Heideggerian terminology: Being cannot be accessed independently of Dasein's *Seinsverständnis*. It is consequently important to recognize the difference between a *prinzipien-theoretisch* and an *evidenz-theoretisch* type of transcendental philosophy.[48] In contrast to (neo-)Kantian transcendental philosophy, which remains merely critical (*prinzipien-theoretisch*), phenomenology has insisted upon the importance of evidence and the first-person perspective.

If we look closer at some of Husserl's scattered remarks about Kant, we will find Husserl faulting Kant for not having had a proper concept of the *a priori*, for operating with a too strong distinction between sensation and understanding, for remaining too metaphysical, for having overlooked a number of foundational issues in his critique of knowledge, for being too oriented towards the natural sciences, for confusing noetic and noematic analyses, and for lacking methodological rigor.[49] When it comes to the difference in method, a telling statement can be found in a manuscript from 1920, where Husserl writes as follows: "Kant's deduction is a paradigmatic example of a transcendental demonstration from above. It remains far

---

[46] Steven Galt Crowell, *Husserl, Heidegger, and the Space of Meaning* (Evanston: Northwestern University Press, 2001), 45, 51, 89.

[47] *Ibid.,* 54, 58.

[48] Jitendranath Mohanty, *The Possibility of Transcendental Philosophy* (Dordrecht: Martinus Nijhoff, 1985), 215.

[49] Husserl, *Erste Philosophie (1923/24). Erster Teil*, 198–99, 235, 282; Husserl, *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie*, 420–21; Husserl, *Logische Untersuchungen II*, 729, 732; Husserl, *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Drittes Buch*, 128; Edmund Husserl, *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie I*, Husserliana III/1–2 (Den Haag: Martinus Nijhoff, 1976), 246; Edmund Husserl, *Die Idee der Phänomenologie. Fünf Vorlesungen*, Husserliana II (Den Haag: Martinus Nijhoff, 1950), 48.

from all phenomenological analyses." Moreover, as he then adds, such a deduction can only be met with disapproval (*Kopfschütteln*) by phenomenologists.[50] Some years later, Husserl expands on this remark, and writes that a transcendental deduction of the *a priori* structures of the world can take two paths. There is the direct way from below, which takes its point of departure in the concrete pre-predicative experience of the world, and then there is the way from above, which takes its point of departure in logic. Needless to say, the first way is the phenomenological way, the second the Kantian way.[51] In one of his longest text on Kant, *Kant and the Idea of Transcendental Philosophy* (*Kant und die Idee der Transzendentalphilosophie*), written and presented in commemoration of Kant's bicentennial in 1924, Husserl writes that transcendental philosophy should be based upon a systematic description and analysis of consciousness in all of its modalities.[52] He then criticizes Kant's method for being regressive-constructive. It lacks an intuitive basis and is unable to provide us with a proper account of consciousness. In fact, as Husserl points out in the conclusion, phenomenology insists upon an in-depth investigation of consciousness, and this demand necessitates an *extension* of Kant's concept of the transcendental. It proves necessary to include the humanities and the manifold of human sociality and culture in the transcendental analysis.[53] This remark is amplified a few years later when Husserl writes that the possibility of a transcendental elucidation of subjectivity and world is lost if one follows the Kantian tradition in interpreting transcendental subjectivity as an isolated ego and thereby ignores the problem of transcendental intersubjectivity.[54]

One commentator has recently argued that Husserl through the 1920s and 1930s "became increasingly wide-reaching, even baroque, in his conception of the transcendental".[55] But rather than calling Husserl's notion of the transcendental baroque, perhaps it would be more to the point to realize that Husserl subjected the very notion of the transcendental to a far-reaching transformation. As I have argued in detail elsewhere, Husserl's later phenomenology is characterized by its attempt to modify the static opposition between the transcendental and the mundane, between the constituting and the constituted.[56] It is, for instance, against this background that one should understand assertions from *Cartesian Meditations* (*Cartesianische Meditationen*) to the effect that the constitution of the world implies a mundanization of the

---

[50] Quoted in Iso Kern, *Husserl und Kant: Eine Untersuchung über Husserls Verhältnis zu Kant und zum Neukantianismus* (Den Haag: Martinus Nijhoff, 1964), 104.

[51] Husserl, *Natur und Geist*, 103, 112.

[52] Husserl, *Erste Philosophie (1923/24). Erster Teil*, 234–35.

[53] *Ibid.,* 282.

[54] Edmund Husserl, *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie. Ergänzungsband. Texte aus dem Nachlass 1934–1937*, Husserliana XXIX (Dordrecht: Kluwer Academic Publishers, 1993), 120

[55] Dermot Moran, "Making Sense: Husserl's Phenomenology as Transcendental Idealism," in Jeff Malpas (ed.), *From Kant to Davidson: Philosophy and the Idea of the Transcendental* (London: Routledge, 2002), 51.

[56] Dan Zahavi, *Husserl und die transzendentale Intersubjektivität,* Phaenomenologica 135 (Dordrecht: Kluwer Academic Publishers, 1996); Zahavi, *Husserl's Phenomenology.*

constituting subject, – that is, that the subject's constitutive experience of the world goes hand in hand with the subject's constitutive experience of its own worldly being.[57] One of the interesting consequences of this idea is that the empirical subject can no longer be regarded as a mere incidental appendix to the transcendental subject, and therefore no longer as something which transcendental philosophy can allow itself to ignore. On the contrary, it is of decisive importance to understand why the transcendental subject as part of the constitutive process must *necessarily* conceive of itself as a worldly entity. In my view, Husserl ultimately argued that the transcendental subject must necessarily conceive of itself as a worldly being if it is to constitute an objective world, since objectivity can only be constituted by a subject which is both *embodied* and *socialized*.[58] Such a claim, however, is in striking contrast with any traditional Kantian understanding of the matter, and eventually Husserl's rethinking of the relation between the transcendental and the empirical made him engage with issues, which a traditional Kantian transcendental philosophy would have denied any transcendental relevance; issues such as *generativity*, *tradition*, *historicity*, and *normality*.

In *The Order of Things* (*Les mots et let choses*, 1966), Foucault argues that phenomenology exemplifies a type of modern discourse that in its investigation of experience seeks to both separate as well as integrate the empirical and the transcendental. It is an investigation of experience that in the face of positivism has tried to restore the lost dimension of the transcendental, but which at the same time has made experience concrete enough to include both body and culture. To Foucault it is quite clear that this modern type of transcendental reflection differs from the Kantian type by taking its point of departure in the paradox of human existence rather than in the existence of natural science. Although Husserl had apparently succeeded in unifying the Cartesian theme of the cogito with the transcendental motif of Kant, the truth is that Husserl was only able to accomplish this union insofar as he changed the very nature of transcendental analysis. It is this transformation that in Foucault's view has resulted in phenomenology's simultaneously promising and threatening proximity to empirical analyses of man.[59]

I think Foucault's diagnosis is correct (though I do not share his subsequent criticism of phenomenology), and I think it holds true not only in the case of Husserl, but also for many post-Husserlian phenomenologists. In the famous preface to *Phenomenology of Perception* (*Phénoménologie de la perception*) Merleau-Ponty calls attention to the simple fact that even 50 years after Husserl's first works, there is still no definitive answer as to what precisely phenomenology is. In fact, the different definitions given all seem to point in different directions. On the one hand, phenomenology is presented as a form of transcendental philosophy. Through

---

[57] Edmund Husserl, *Cartesianische Meditationen und Pariser Vorträge*, Husserliana I (Den Haag: Martinus Nijhoff, 1950), 130.

[58] Husserl, *Cartesianische Meditationen und Pariser Vorträge*, 130; Husserl, *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Drittes Buch*, 128; Husserl, *Die Idee der Phänomenologie*, 162.

[59] Michel Foucault, *Les mot et les choses* (Paris: Gallimard, 1966), 331–36.

reflection it seeks to unearth the conditions of possibility for experience, and it suspends the assumptions we make in the natural, pre-philosophical, attitude precisely in order to investigate them critically. But at the same time, phenomenology admits that reflection must always take its point of departure in an already existing world-relation, and that the philosophical task is to articulate the full significance of this immediate and direct contact with the world. Phenomenology is a transcendental philosophy, but its point of departure remains factual existence. In this sense, phenomenology is also a thinking of facticity. As Merleau-Ponty then remarks, it could be tempting to overcome this apparent contradiction by simply distinguishing between Husserl's transcendental phenomenology, which has often been interpreted as the attempt to thematize the pure and immutable conditions of knowledge, and Heidegger's hermeneutical and existential phenomenology, which has often been interpreted as an attempt to point to the historical and practical contextuality of knowledge. But Merleau-Ponty then goes on to reject this solution as being much too simple. First of all, the "contrast" in question can be found internally in Husserl's writings, and secondly, and more importantly, we are not in any way faced with real contrasts or alternatives, but with complementary aspects that phenomenology must seek to integrate.[60] To put it differently, one of the defining features of phenomenological thought has been its attempt to make the co-existence between the transcendental and the empirical perspective less paradoxical. Rather than conceiving of the two as mutually incompatible, they are intertwined and complementary perspectives.[61]

The fact that transcendental phenomenology operates with an enlarged notion of the transcendental, the fact that it includes topics such as embodiment and intersubjectivity into its transcendental analysis, gives it an advantage in comparison with a more traditional Kantian type of transcendental philosophy. But, of course, it would also be fair to say that this transformation generates new problems of its own.

13.4 Conclusion

Phenomenologists have insisted upon the difference between an empirical and a transcendental analysis of consciousness. This is not to say that philosophy should ignore empirical science and evidence, but to paraphrase Putnam, it is entirely possible to insist that philosophy needs to be informed by the best available scientific

---

[60] Maurice Merleau-Ponty, *Phénoménologie de la perception* (Paris: Gallimard, 1945), i–ii.

[61] In a recent article, Steven Crowell has argued that the issue of facticity and the emphasis on the situated or concrete nature of subjectivity, rather than rendering transcendental philosophy impossible, can actually illuminate the specific character of transcendental phenomenology. In fact, Heidegger's introduction of the notion of *Dasein* is precisely meant to undermine the claim that there can be no philosophical but only an empirical-psychological inquiry into concrete subjectivity, see Steven Galt Crowell, "Facticity and Transcendental Philosophy," in Jeff Malpas (ed.), *From Kant to Davidson: Philosophy and the Idea of the Transcendental* (London: Routledge, 2002), 100, 105.

knowledge, while at the same time insisting that philosophical and scientific questions differ fundamentally.[62]

Although it is important to encourage the exchange between philosophy and empirical science, the possibility of a fruitful cooperation between the two should not make us overlook their difference. Husserl for his part would insist that although there might be a considerable overlap when it comes to the content of transcendental and empirical science – he even speaks of a parallelism between the two – their difference is precisely not a difference in content, but in attitude. Phenomenology and psychology both investigate consciousness, but they do so with quite different agendas in mind, and it would make just as little sense to identify the phenomenologists' transcendental analysis of consciousness with a psychological investigation as it would be to identify, say, Wittgenstein's or Davidson's language philosophical investigations with phonetic or linguistic analyses.

Husserl has frequently been accused of being a foundationalist. To some extent this is correct. Husserl is a transcendental philosopher, and he would insist that transcendental phenomenology investigates the condition of the possibility for experience, meaning, and manifestation, and thereby also the very framework of intelligibility that conditions every scientific inquiry. However, Husserl also quite explicitly denies that transcendental subjectivity could ever serve as the starting point for a transcendental deduction.[63] Phenomenology is not a *deductive* discipline, but a *descriptive* discipline.[64] To put it differently, the truths that transcendental phenomenology might uncover do not make up a foundation which the contents of the positive sciences could be deduced from. Moreover, Husserl did not conceive of his own transcendental analysis as a conclusive, final analysis. It is an exploration of a field, which, in an absolute sense, is unavoidable (*unhintergehbar*). But the analysis of this field can always be refined, deepened, and improved. Thus, for Husserl the full and conclusive truth about the transcendental dimension remains a regulative ideal. Philosophy as a science based on ultimate justification is an idea which can only be realized in an infinite historical process.[65]

One way to understand Husserl's foundationalism might consequently be to see it as a way of emphasizing the difference between the empirical and the philosophical stance, between the mundane and the transcendental attitude. It is by insisting upon this difference that Husserl can maintain that philosophy does offer something cognitively distinctive to the scientific enterprise.[66]

---

[62] Hilary Putnam, *Renewing Philosophy* (Cambridge: Harvard University Press, 1992), 34.

[63] Husserl, *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie*, 193; Husserl, *Cartesianische Meditationen und Pariser Vorträge*, 63.

[64] Husserl, *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie I*, 158.

[65] Husserl, *Erste Philosophie (1923/24). Zweiter Teil*, 186; Husserl, *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie*, 439.

Chapter 14
Phenomenological Responses
to Gestalt Psychology

Sara Heinämaa

Abstract The chapter studies the reception that Gestalt psychological theories were given by phenomenologists in Germany and France in the first half of the twentieth century. The aim is to study, in particular, the reactions of two phenomenologists, Edmund Husserl and Maurice Merleau-Ponty. The chapter focuses on these two thinkers in order to explicate the main idea of the phenomenological-transcendental critique of psychological theories. The interpretative claim is that Merleau-Ponty followed Husserl in defining phenomenological philosophy by its radical task in providing a transcendental basis for all experience and knowledge. He thus came to argue that psychological theories, Gestalt theories included, must be submitted to a phenomenological-transcendental critique. Despite their apparent differences, Merleau-Ponty and Husserl agreed that no empirical or wordly knowledge – psychological, anthropological or natural scientific – can overrule radical philosophical reflections in the grounding of the positive sciences. Before entering into Gestalt-theoretical and phenomenological sources, the chapter briefly discusses the historical relations between the two fields of research. The connections are to be found in common conceptions of parts and wholes, both approaches being influenced by Brentano's distinctions between different kinds of parts. The disparity concerns the role of consciousness in the institution and establishment of meaning.

Keywords Gestalt · parts and wholes · sensation · genetic phenomenology · Merleau-Ponty · Husserl · Brentano

Edmund Husserl is well-known for his detailed arguments against different forms of naturalism and psychologism, forms which all aim at establishing a foundation for knowledge and the sciences through factual inquiries into the natural world (or into the psyche as part of the natural world). Husserl's main critique against such attempts was that they fail to differentiate between forms of generality: the factual generality of empirical judgments is confused with the exceptionless universality of eidetic and transcendental claims. In his early work *Logical Investigations* (*Logische*

S. Heinämaa
Helsinki Collegium for Advanced Studies, The Academy of Finland, University of Helsinki

Untersuchungen 1900–1901), Husserl resolutely attacked both psychologistic and biologistic philosophies of logic and knowledge. In a long article, *Philosophy as Rigorous Science* (*Philosophie als strenge Wissenschaft* 1911), he also questioned historicist philosophies of knowledge and argued that they too suffer from the same principal form of confusion as psychological and biological naturalism. The mature work, *Ideas to a Pure Phenomenology and Phenomenological Philosophy* (*Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie* 1913), brought these perspectives together and argued that only a transcendental inquiry into pure consciousness, its necessary eidetic structures, can provide a foundation for the sciences. This is the task of transcendental-phenomenological philosophy.

14.1 A Dispute Over Gestalt Psychology: Husserl and Gurwitsch

In 1931, Husserl published a preface to the English translation of his *Ideas*.[1] In this text, he explicitly included Gestalt psychology among the fields which suffer from the confusion of transcendental-eidetic and mundane (wordly) knowledge. Husserl saw Gestalt theories as the latest development of the Lockean tradition which had been misled by the model of the objective sciences of nature.[2] The main problem with this tradition was that it did not grasp the radicalism of Descartes' innovation of "ego cogito" but treated the realm of consciousness as a specific area of natural reality.[3] Thus, the basic elements of consciousness – the data of inner

---

[1] This text is known by the title "Nachwort" and it is included in *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Drittes Buch: Die Phänomenologie und die Fundamente der Wissenschaften*, Husserliana V (Den Haag: Martinus Nijhoff, 1971), 133–62. An English translation is included as "Epilogue" in *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, Second Book: Studies in the Phenomenology of Constitution*, trans. R. Rojcewicz and A. Schuwer (Dordrecht, Boston, London: Kluwer Academic Publishers, 1989), 405–30; and as "Preface to the English Translation," in *Ideas: General Introduction to Pure Phenomenology*, trans. W.R. Boyce Gibson (New York, London: Collier, 1931), 5–22.

[2] *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Drittes Buch*, 157; *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, Second Book*, 424; cf. Edmund Husserl, *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie. Eine Einleitung in die phänomenologische Philosophie*, Husserliana VI (Den Haag: Martinus Nijhoff, 1950), 87–8, 344–45; *The Crisis of European Sciences and Transcendental Phenomenology*, trans. David Carr (Evanston: Northwestern University Press, 1970), 85, 297.

[3] In the late work, Husserl states quite vehemently: "Locke senses nothing of the depths of the Cartesian *epoche* and of the reduction to the ego. He simply takes over the ego as soul, with becomes acquainted, in the self-evidence of self-experience, with its inner states, acts, or capacities." (*Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie*, 86; *The Crisis of European Sciences*, 84). Husserl continues by arguing that Locke's inability to grasp the structure of intentionality, implied by Descartes' *ego-cogito*, is especially disastrous. This neglect results in the separation of conscious states and processes from their original and essential connections with the world. "The inner realm" is left receptive merely to causal effects, which are assumed to issue from the world as it is known and believed to be. Thus, we end up with the problematic picture of the mind as "an inner theater". Husserl writes: "Especially portentous for future psychology

experience or sense-data – were distinguished and classified in an analogous way as spatial-temporal entities, and then inserted into combinatory and causal relations. The basic presumption of naturalism was shared, in Husserl's view, both by the British empiricist and by his contemporary Gestalt theorists. Husserl argued:

> Evidently and in principle, it makes no difference in this regard whether one lets the psychic data be blown into aggregates "atomistically," like shifting heaps of sand [...] or whether they are considered parts of wholes which, by necessity, [...] can behave individually only as such parts within a whole [...] In other words, atomistic psychology, as well as Gestalt psychology, both retain the sense and the principle of "naturalism" [...] or "sensualism," as it can also be named if we recall the use of the term "inner sense."[4]

Thus the main trouble with Gestalt theory was the same as with all naturalistic theories of the mind:

> For those who live in the habits of thought prevailing in the science of nature it seems to be quite obvious that purely psychic being, or psychic life, is to be considered a course of events similar to natural ones, occurring in the quasi-space of consciousness.[5]

The Lockean tradition is often characterized as "phenomenological", Husserl remarked, but it must be strictly distinguished from phenomenology proper because its field of study, its aims, and its methods are crucially different from those of the phenomenologists. In the Lockean tradition, consciousness and mental states are studied as real entities or events of the inner world having real causal and combinatory relations. Phenomenology, on the contrary, aims at suspending all positing of the world and its reality, and studies consciousness as the source of all the sense that reality has and can have. The relations that are thus disclosed and illuminated are not real or causal relations, but intentional and constitutional relations pertaining to pure consciousness.

Thus, the research object of the phenomenologist is not mind or consciousness as a worldly or mundane (spatial-temporal) object, among other objects, psychic and physical. When studied through phenomenological methods, consciousness is not given as a part of the world but as our fundamental intentional relation to the world,

---

and theory of knowledge is the fact that Locke makes no use of the Cartesian first introduction of the *cogitatio* as *cogitatio* of *cogitata* – that is, intentionality; he does not recognize it as a subject of investigation [...] He is blind to the whole distinction. The soul is something self-contained and real by itself, as is a body; in naive naturalism the soul is now taken to be like an isolated space, like a writing tablet [...] on which psychic data come and go. [...] Of course one speaks quite unavoidably, [even] in the Lockean terminology, of perceptions, representations 'of' things, or of believing 'in something,' willing 'something,' and the like. But no consideration is given to the fact that in the perceptions, *in the conscious experiences themselves* [*Bewusstseinserlebnissen*], that of which we are conscious is included *as such*; [no consideration is given to the fact] that the perception is *in itself* a perception *of* something, of 'this tree' " (*ibid.*, 87–8; 85, translation modified).

[4] *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Drittes Buch*, 156–57; *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, Second Book*, 423–24.

[5] *Ibid.*, 156; 423.

an opening onto it. In the preface, Husserl characterizes the new insight of his own approach as follows:

> What is essentially new, broken open in transcendentally oriented phenomenology [...] is the insight that a concrete description of the sphere of consciousness as a self-enclosed sphere of intentionality [...] has a total different sense than descriptions of nature, thus than the exemplary descriptions in the descriptive natural sciences.[6]

Husserl emphasizes that consciousness is given concretely only as a sphere of intentionality. Thus, what needs to be studied and described are not only intentional objects as correlates of intentional acts, but also the genetic constitution of both types of unity in the temporal flow of lived experiences.[7]

Husserl explains again and again that phenomenology is not a new philosophical theory or a new development in one of the traditional philosophical subfields, i.e. philosophy of mind or philosophy of cognition.[8] Rather it is, and must be understood as, a new science with its own peculiar aims and methods and its own field of study: pure experience and the constitutive work of consciousness. As such phenomenology is a universal and founding science of all other sciences, that is, the first philosophy in Descartes' sense.[9]

In 1932, Husserl's young successor and critic Aron Gurwisch published a review of the preface that Husserl wrote to his *Ideas*.[10] Gurwisch had worked to combine phenomenological and Gestalt-theoretical approaches and saw Husserl's opposition as misguided and prejudiced. His main argument was that Husserl neglected important points of convergence between the two approaches because he mistakenly interpreted the main innovation of Gestalt theory to be in its holistic account of sensation (vs. the atomistic theories of classical and contemporary empiricists). According to Gurwitsch, this innovation was secondary and resulted from a more profound methodological insight: the Gestalt theorists abandoned the idea

---

[6] *Ibid.*, 157; 424.

[7] *Ibid.*, 157; 424–25.

[8] In this context, Husserl also answers the critics who discard phenomenology as just another form of traditional idealism. He argues that traditional realism as well as traditional idealism, its opposite alternative, must both be rejected in so far as they take an area of reality – the material or the mental – as their foundation: "I may not here neglect, however, to declare expressly that I retract nothing whatsoever as regards transcendental-phenomenological idealism and that I still consider, as I did before, every form of the usual philosophical realism nonsensical in principle, no less so than that idealism which it sets itself up against in its argumentation and which it 'refutes'" (Husserl, *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Drittes Buch*, 152; *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, Second Book*, 418; cf. *Cartesianische Meditationen und Pariser Vorträge*, Husserliana I [Den Haag: Martinus Nijhoff, 1950], 116–19; *Cartesian Meditations: An Introduction to Phenomenology*, trans. Dorion Cairns [Dordrecht, Boston, Lancaster: Martinus Nijhoff, 1960], 83–7.)

[9] Cf. Husserl, *Cartesianische Meditationen und Pariser Vorträge*, 182–83; *Cartesian Meditations*, 156–57.

[10] Aron Gurwitsch, "Critical Study of Husserl's *Nachwort*," trans. Frederick Kersten, in *Studies in Phenomenology and Psychology* (Evanston: Northwestern University Press, 1966), 107–15.

of isolated sense-data because they aimed at liberating the description of experience from all conceptions concerning its correlation with external stimuli (or "pregiven objectivity").[11] In a letter, Gurwitsch describes this Gestalt theoretical insight in reference to Wolfgang Köhler's work as follows:

> In an article of 1913, "Über unbermerkte Empfindungen und Urteilstäuschungen," Koehler showed that modern psychology, particularly in the nineteenth century, proceeds on the taken-for-granted and hardly ever explicitly formulated assumption that *sense data*, as the ultimate elements of conscious life, *depend exclusively and exhaustively on local stimulation*: when a sense organ is stimulated in the same way, the same sensation is bound to arise. Koehler called this assumption the "constancy hypothesis." In it the logicohistorical continuity of modern physics and modern psychology is apparent: the latter relies upon the former and avails itself of its results.[12]

Gurwitsch saw the Gestalt-theoretical rejection of the constancy hypothesis as an incipient or embryonic form of Husserl's phenomenological *epochē*. The common aim of both approaches was to describe the elements of perceptual and sensuous experience as they present themselves, without any reference to objectivities belonging to or acquired from other areas of experience, such as physical or physiological explanation and theorization. Thus, Gurwitsch argued that Gestalt theorists, in a very similar way as phenomenologists, proceed by suspending prejudices of "pregiven objectivities" and a "pregiven world".

Moreover, Gurwitsch suggested that the Gestalt theorist were more successful, or more uncompromising, in their suspensions than Husserl was. He argued that Gestalt-psychological results challenged Husserl's phenomenological theory of intentionality. This argument was based on Gurwitsch's conviction that Husserl's distinction between sense-giving or meaning-bestowing acts and the so-called *hyletic data* (which is articulated by such acts) implies the idea of senseless matter without any articulation or structure.[13] Gestalt-theoretical results proved that such a notion of sensuous matter was a theoretical postulate, coined for the purposes of explanation, with no experiential justification.

Gurwitsch did not intend to abandon or reject Husserl's phenomenology; rather he undertook to improve its concepts insofar as he found them to conflict with Gestalt-theoretical results. This led, however, to a fundamental problematization of Husserl's concepts of intentionality and constitutional analysis:

> Among those doctrines is Husserl's dualistic theory of perception as it manifests itself in his distinction between hyletic data and apperceptive or functional characters [i.e. intentional characters]. We have advocated that the notion of hyletic data be relinquished altogether. Its relinquishment would entail a revision of the theory of intentionality. Intentionality can no longer be generally defined as bestowal of sense, for, in the last analysis, such a definition

---

[11] Gurwitsch presents this argument clearly in his review of Husserl's preface, "Critical Study," 113–14.

[12] Cited in Lester Embree, "Biographical Sketch of Aron Gurwitsch," in Lester Embree (ed.), *Life-World and Consciousness* (Evanston, Illinois: Northwestern University Press, 1972), viii.

[13] Aron Gurwitsch, *Field of Consciousness* (Pittsburg: Duquesne University Press, 1964), 28, 79, 90–91; cf. Lester Embree, "Gestalt Psychology," in Lester Embree, Elisabeth A. Behnke, David Carr et al. (eds.), *Encyclopedia of Phenomenology* (Dordrecht: Kluwer, 1997), 279.

implies that, at the basis of the structural hierarchy of intentional functions and accomplishments, there is a bestowal of sense on hyletic data which in themselves are devoid of it [sense]. [. . .] Allowance for certain trends in contemporary psychology, especially Gestalt psychology, has lead us to the discussed reexaminations, revisions, and even departures.[14]

Gurwitsch argued for this view already in 1928 in his doctoral work, *Phänomenologie der Thematik und des reinen Ich: Studien über Beziehungen von Gestalttheorie und Phänomenologie*.[15] He continued to develop his argument despite the fact that he never succeeded in convincing Husserl about the putative connection between Husserl's account of intentionality (its concept of *hyletic data*) and the problematic idea of primary structureless matter.[16] On the contrary, Husserl vehemently questioned the self-understanding of Gestalt psychologists, arguing that this psychology, like any other psychology, necessarily rests on certain validities and cannot account for their constitution.[17] Providing such an account was for him the true task of philosophy.

I will come back to this divergence later, when I discuss the philosophical treatment that Maurice Merleau-Ponty gave to the Gestalt-psychological discourse and its results, most importantly the idea of the *body schema*. Before entering more deeply into the phenomenological and Gestalt-psychological sources, I will shortly discuss the historical relations between the two modes of inquiry.

14.2 Brentano's Teaching: Separable and Dependent Parts

The relationship between phenomenology and Gestalt psychology is not just critical. The two fields share a common historical source: both movements have roots in Franz Brentano's philosophy of consciousness and his related metaphysical theory of parts and wholes. Both emerged, by affirmation and critique, from the new "descriptive psychology" that Brentano introduced and developed in his teachings

---

[14] Aron Gurwitsch, "Edmund Husserl's Conception of Phenomenological Psychology," in *Phenomenology and the Theory of Science*, ed. Lester Embree (Evanston, Illinois: Northwestern University Press, 1974), 111–12. Several American phenomenologists have been influenced, directly or indirectly, by Gurwitsch critical interpretation of Husserl's phenomenology, and many also read Merleau-Ponty's works in a Gurwitschian framework. See, for example, Martin Dillon, "Gestalt Theory and Merleau-Ponty's Concept of Intentionality," *Man and World* 4 (1971), 436–59.

[15] Aron Gurwitsch, "Phänomenologie der Thematik und des reinen Ich: Studien über Beziehungen von Gestalttheorie und Phänomenologie," *Psychologische Forschung*, 12 (1929), 279–383. An English version of this text is included in the collection *Studies in Phenomenology and Psychology* (Evanston, Illinois Northwestern University Press, 1966), 175–286.

[16] Embree, "Biographical Sketch," xx. For Husserl's critique of the Gurwitschian reading of his theory of intentionality, see his *Formale and transzendentale Logik: Versuch einer Kritik der logischen Vernunft*, Husserliana XVII (Den Haag: Martinus Nijhoff, 1974 [1929]), 252–53; cf. Robert Sokolowski, *The Formation of Husserl's Concept of Consciousness* (The Hague: Martinus Nijhoff, 1964), 177–81.

[17] See, e.g., Husserl, *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie*, 344–45; *The Crisis of European Sciences*, 296–97.

and publications at the end of the nineteenth century.[18] The line of inheritance is direct and concrete: Brentano was the teacher of Husserl as well as Christian von Ehrenfels, the first Gestalt theorist, so called.

In 1890, Christian von Ehrenfels published an article in philosophical psychology, titled *On "Gestalt qualities"* (*Über "Gestaltqualitäten"*). This work has been commented on by most later Gestaltists, and it provides the common reference point for the different Gestalt-theoretical schools, such as the Berlin School of Max Wertheimer, and the Graz School founded by Alexius Meinong.[19] Ten years later, in 1900–1901, Husserl published his *Logical Investigations* (*Logische Untersuchungen*) which became the cornerstone of phenomenological philosophy and which includes an elaborated critique of all forms of naturalism, Brentano's descriptive psychology included. A common conceptual source for both these works was in Brentano's theory of parts and wholes.[20]

In his lectures on *Descriptive Psychology* (*Deskriptive Psychologie* 1887–1891) and in his later essays on *The Theory of Categories* (*Kategorienlehre* 1907, 1917, 1933), Brentano laid down the foundations for a metaphysical theory of parts and wholes. What we find in his lectures and other writings is not a fully developed systematic mereology, but methodological reflections which provide a series of important distinctions. These distinctions were groundbreaking, and were reinterpreted and developed by Husserl (in his third Logical Investigation).

Brentano developed his concepts of parts and wholes in order to make sense of his distinction between *pure descriptive psychology* (or *psychognosy*) and *genetic psychology*. The former aims at disclosing the necessary elements and structures of human consciousness, i.e. its basic mental components and their ways of connecting. The latter provides causal explanations of the physiological and physical conditions under which particular or specific psychic phenomena occur in human minds. The explanations of genetic psychology are methodologically secondary and dependent on the descriptive results in the sense that they presuppose the units and structures which only pure psychology is able to disclose and explicate.[21]

Brentano's main thesis was that consciousness is a unified whole with different kinds of parts and a rich inner structure. Thus, he argued for the Aristotelian and Cartesian conception of the unity of mental phenomena and rejected the Humean

---

[18] The tradition of Gestalt theory was also influenced by developments in physics, most importantly by the concepts of field theory. See e.g., Kurt Koffka, *Principles of Gestalt Psychology* (London: Routledge and Kegan Paul Ltd., 1938 [1935]), 41–2, 53.

[19] Barry Smith even states: "Almost all of the theoretical and conceptual issues which came subsequently to be associated with the Gestalt idea are treated at some point of the work [Ehrenfels' essay], at least in passing" (Barry Smith, "Gestalt Theory: An Essay in Philosophy," *The Foundations of Gestalt Theory* [München, Wien: Philosophia Verlag, 1988], 15).

[20] The teacher-pupil relations are more complex than this. For example, Max Wertheimer, the leading figure of the Berlin School, regularly attended Ehrenfels' lectures, and Kurt Koffka, another member of the same School, was a disciple of Husserl. For a full account of these influence relations, see, Smith, "Gestalt Theory."

[21] Franz Brentano, *Descriptive Psychology*, trans. and ed. Benito Müller (London and New York: Routledge, 1995), 3–4, cf. 78, 156, 165–66.

"bundle theory" which claims, or suggests, that consciousness is nothing but a collection of mental events or occurrences which have no internal relations but are merely bound together by associative and causal links. Despite his indebtedness to Aristotle and Descartes, Brentano praised Hume's insight into the complexity and composition of mental phenomena: "However unfortunately Hume's comparison of our consciousness to a bundle, he is undoubtedly correct on one point: Our consciousness does not present itself to our inner perception as something simple, but it shows as being composed of many parts."[22]

According to Brentano, the main problem with traditional approaches to consciousness was that they lacked adequate conceptual distinctions to account for the internal relations and structures of conscious states and processes. Most modern theorists operated on the assumption that the parts of consciousness are spatial or quasi-spatial parts (*partes extra partes*), succeeding one another or lingering side-by-side in the internal realm. Accordingly, the ultimate constituents of consciousness were conceived of as spiritual atoms. Brentano rejected the idea of spatial or quasi-spatial separation and proceeded by developing an alternative idea of separation. He defined two mental states as *separable* (*ablösbar*) in so far as one of them can continue to exist when the other has ceased to be.[23] For example, when attending a concert, I can see the orchestra and simultaneously hear it playing. My two simultaneous perceptions are mutually separable in Brentano's terms, because I can hear the music even when I close my eyes, and conversely, covering my ears does not prevent me from seeing the musicians playing their instruments. Thus, we can say that seeing and hearing the orchestra are two mutually separable parts of one and the same complex conscious state.

Hearing and seeing exemplify *mutually* separable parts, but Brentano points out that consciousness also includes *one-sidedly* separable parts. Such a part-part relationship holds, for example, between the activities of seeing and attending: I cannot attend to the sound of the violin without hearing it, but I can hear the sound, Brentano argues, without attending to it.[24]

Not all the parts of consciousness are separable, however, neither mutually nor one-sidedly separable. On the contrary, most psychic parts are interdependent and belong to consciousness only as bound together. An example of such a bondage is the connection between sensed qualities, such as color and spatial determination. Neither the color nor the spatial determinants of a spot can be changed, Brentano argues, without removing the spot and replacing it by another spot.[25] Brentano calls such interdependent parts "distinctional" (*distiktionell*), and argues that consciousness is rich in them. Most importantly, he argues that the *intentional* structure which characterizes all conscious states includes two principal parts which are

---

[22] Brentano, *ibid.*, 15, cf. 83.

[23] Brentano, *ibid.*, 15.

[24] Brentano, *ibid.*, 15, 25–6.

[25] A less controversial example of distinctional relations would be the relations which hold between the hue and the saturation in a color sample.

mutually distinctional: the *intentional act* and the *intentional object*. These cannot exist independently of one another, but can only be differentiated by concepts.[26]

With these conceptual tools Brentano was able to question the prevailing understanding of the unity and coherence of the mind. He attacked the modern notion according to which the self is a non-material entity which supports its mental states in the same way as a bridge supports the various vehicles that pass over it. According to this picture, mental states would be occurrences which pass the mind's unmovable eye, or like appendices attached to the mind.

As is well known, Brentano returned to pre-modern sources and reinterpreted Aristotle's psychological theory according to which the perceiving soul is able to include or contain in itself the sensible forms of things without incorporating their matter. Brentano revised the Aristotelian idea of inclusion with his concepts of parts and wholes. He argued that the perceiving mind should not be conceived as a whole with two separable parts – the thing which is able to perceive and the perception that it occasionally entertains. Rather, the whole in question – the perceiving mind – is a result of an enrichment or augmentation of its proper (non-identical) part. In *The Theory of Categories*, Brentano explains:

> Among the entities that have parts, there are some whose whole is not composed of a multiplicity of parts; it appears much rather as an enrichment of a part, though not as a result of the addition of a second part. One example of such an entity is a thinking soul. It ceases to think and yet remains the same soul. But when it starts to think again no second thing is added to that entity which is the soul. What we have here, then, is not like what we have when one stone is laid alongside another or when we double the size of the body.[27]

Thus, the modified soul – the thinking soul, the perceiving soul, or the desiring soul – is not a whole composed of two separable parts. The non-modified soul is included in the modified one as a separable part; but the modification is not a

---

[26] Brentano, *ibid.*, 24ff, cf. to his *Psychology from an Empirical Standpoint*, trans. Antos C. Rancurello, P.B. Terrell, and Linda L. McAlister (London: Routledge, 1995 [1874]), 88–9. Brentano's concept of distinctional parts covers several different types or classes of psychic components. So the parts of the intentional relation – the intentional pair – are one type of distinctional parts. Other distinctional parts are of the types: (i) mutually pervading (or "concrescent" or concrete) parts, e.g. the parts of truth-claims, such as the quality of affirming and the modality of apodicticity, (ii) (one-sidedly) logical parts, e.g. "experiencing is a logical part of seeing, and seeing is a logical part of seeing-red" (Brentano, *Descriptive*, 178), (iii) the parts of psychical "di-energy". Moreover, some parts are properly distinctional and others distinctional merely through modification. For these distinctions, see, Brentano, *ibid.*, 22ff., 83ff.

Brentano also famously argued that each conscious state is by necessity directed at itself so that it has itself as a secondary object: "The fact that there is no consciousness without any intentional relation at all is as certain as the fact that, apart from the object upon which it is primarily directed, consciousness has, on the side, itself as an object. This is, in an essential way, part of the nature of every psychical act" (Brentano, *ibid.*, 26). He attributed both claims to Aristotle, referring to his *Metaphysics* 1021a30 and *De anima* 425b (Brentano, *ibid.*, 24–6, 180).

[27] Franz Brentano, *Kategorienlehre*, ed. A. Kastil (Leibzig: Meiner, 1933), 47. For a detailed account of Brentano's theory of parts and wholes, see, Barry Smith, *Austrian Philosophy: The Legacy of Franz Brentano* (Chicago, Illinois: Open Court Publishing Company, 1994); cf. also Eduardo Fugali's chapter in this volume.

separable part; it is an extension or an augmentation of the non-modified soul. In *Descriptive Psychology*, Brentano states:

> Everything psychical which we apprehend is composed. It is an accident which includes the substance of the soul, or a plurality of accidents of the same substance, each of which contains the substance.[28]

Both Husserl and Carl Stumpf, Husserl's colleague and friend, revised and developed further the basic distinctions of Brentano's mereological reflections, and based their respective discussions on perception and experience on this conceptual foundation. Husserl dedicates the third Logical Investigation to the explication of the fundamental distinctions between dependent vs. independent parts, founded vs. founding parts, and concrete wholes vs. abstract parts.[29] Ehrenfels too was influenced directly by Brentano, but equally importantly by Stumpf. The basic problem that he tackled, however, – the problem of the emergence of Gestalt qualities – comes from the positivist Ernst Mach, as we will see.

Thus, classical Husserlian phenomenology and Gestalt theory can historically be described as two rivers taking different directions – methodologically and thematically – from the stream of Brentano's philosophy of mind.[30] However, the metaphor of rivers is historically correct only if we keep in mind that both movements also had other important background factors and were not simply developments within the Brentanian School.

Systematically, the relation is more problematic. Husserl argued that his phenomenology questions not only his contemporary Gestalt theories, but also all variants of descriptive psychology, the Brentanian as well as the Diltheyan versions.[31] In the first edition of *The Logical Investigations*, he still used the term "descriptive psychology" for his own approach, but in the revised edition published in 1913, he had already abandoned the term and separated his approach clearly from descriptive

---

[28] Brentano, *Descriptive*, 168.

[29] For an illuminating account of Husserl's distinctions, see Robert Sokolowski, "Logic of Parts and Wholes," *Philosophy and Phenomenological Research* 28 (1968), 537–53.

[30] Cf. Michael Dummett, *Origins of Analytic Philosophy* (London: Duckworth, 1993), 26.

[31] See, for example, Edmund Husserl, *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Zweites Buch: Phänomenologische Untersuchungen zur Konstitution*, Husserliana IV (Den Haag: Martinus Nijhoff, 1952), 172–73, 312–15, 393; *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, Second Book: Studies in the Phenomenology of Constitution*, trans. R. Rojcewicz and A. Schuwer (Dordrecht, Boston, London: Kluwer Academic Publishers, 1989), 181–82, 325–28, 402; cf. *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie*, 225–26, 248–49; *The Crisis*, 222–23, 245–46; see also Husserl's discussion of pure psychology in his *Encyclopaedia* article, "Phenomenology," trans. C.V. Salmon, *The Encyclopaedia Britannica*, 14th ed. vol. 17 (1929), 699–702; this article is included in several collections, for example, in Roderick M. Chisholm (ed.), *Realism and the Background of Phenomenology* (IL: The Free Press of Glencoe, 1960), 118–128; Edmund Husserl, *Psychological and Transcendental Phenomenology and the Confrontation with Heidegger (1927 1931). Collected Works: Volume 6* , trans. and ed. Thomas Sheehan and Richard E. Palmer, (Dordrecht: Kluwer, 1997); Joseph J. Kockelmans (ed.), *Edmund Husserl's Phenomenology* (West Lafayette, Indiana: Purdue University Press, 1994).

psychologies: "I felt its defects [the defects of the first edition of *Logical Investigations*] immediately after its appearance, and also found immediate occasion to object to my misleading account of phenomenology as a descriptive psychology." Referring to a review article that he published in 1903,[32] Husserl continued: "Some of the main points of principle are there briefly but sharply characterized. The psychological description performed in inner experience appears as put on a level with the description of external events in nature performed in external description, but it is, on the other hand *opposed* to phenomenological description, from which all transcendent interpretations of immanent data, even those of psychical acts and states of a real ego, are entirely excluded."[33]

Husserl's distinction between transcendental and mundane investigations of experience is crucial for the understanding of all later interchange between phenomenologists and psychologists. Maurice Merleau-Ponty's dual reaction to Gestalt-psychological approaches must also be interpreted within this methodological framework. So in order to capture the solid content of his constructive and critical, positive and negative, references to Gestalt psychology, we must proceed in two steps: the next part of this chapter introduces the Gestalt theoretical tradition and its main results; the last part relates this discourse to Merleau-Ponty's phenomenological-genetic account of the fundamental form of the body schema.

14.3 Two Conceptions of Gestalt Forms: From Ehrenfels to Wertheimer

Christian van Ehrenfels applied Brentano's concepts in his analysis of perception and developed an influential account of perceptual objects and their structural features. The paradigmatic example that he discussed was spatial shape, but his aim was to extend his account to cover temporal structures, such as melodies and rhythmic figures, as well. He argued that all perceptions of such objects involve presentations of Gestalt forms or Gestalt qualities, as he called them. Such forms are not peculiar to sensory registers, but can also combine elements from different sensory fields. Moreover, higher order forms can be composed of lower ones. Thus, Gestalt forms can be perceived which involve visual, tactile as well as aural elements; and there can be forms that have other forms as their immediate constituents instead of elementary contents.

The basic concepts that Ehrenfels employed in formulating his problem and his solution were Brentano's notions of *presentation*[34] (*Vorstellung*) and *separability*

---

[32] This review article was published in *Archiv für systematische Philosophie*.

[33] Edmund Husserl, *Logische Untersuchungen. Erster Band: Prolegomena zur reinen Logik*, Husserliana XVIII (Den Haag: Martinus Nijhoff, 1975 [1900, 1913]), 12–13; *Logical Investigations, Volume I*, trans. J.N. Findlay (London and New York: Routledge, 1982), 6.

[34] The function of presentation is to bring something before the mind, without explicitly affirming or denying this something. In Brentano's words: "We speak of presentation whenever something appears to us" (Brentano, *Psychology*, 198). This "something" can be of different sorts, e.g. a color

or *independence*. Ehrenfels used these concepts to reformulate and develop further Ernst Mach's argument to the effect that spatial shapes and temporal forms are given already in sense-perception and are not added by any other activity or faculty of the mind.[35]

Ehrenfels defined Gestalt as "a positive content of presentation bound up in consciousness with the presence of complexes of mutually separable (i.e. independently presentable) elements."[36] The complex of presentations which is necessary for the existence of a given Gestalt he called its "foundation" (*Grundlage*). So in Ehrenfels understanding, the Gestalt form is dependent in its existence on the presentational elements that constitute its foundation. But these constituents, for their part, are independent of the Gestalt, and can appear separately.

Ehrenfels, following Brentano and Mach, asked himself:

> [. . .] what precisely the given presentational formations (spatial shapes and melodies) in themselves are. Is a melody (i) a mere sum (*Zusammenfassung*) of elements, or (ii) something novel in relation to this sum, something that certainly goes hand in hand with, but is distinguishable from, the sum of elements?[37]

Ehrenfels set about defending the view that temporal or spatial figures, as experienced, are not merely the sums of their constituent parts, but include additional Gestalt qualities.[38] According to him, this holds for spatial figures as well as for temporal figures, the elements of which are not simultaneously presented. A perception of a melody, for example, includes simple (memory) presentations of the individual tones, but in addition to them and to their sum, also a presentation of

---

or an ideal object such as a triangle. Thus, there are as many kinds of presentations as there are mental contents or objects: presenting a color, presenting a triangle, presenting a general concept. Understood in this way, presentation is an act of mental seeing or entertaining of something, an individual, a concept, or a complex relation.

Brentano took for granted a classification of mental acts into three different types or modes. He spoke of three different basic classes (*Grundklassen*) of mental acts. This classification is traditional, and Brentano found it in Descartes' *Meditations* (the third Meditation) as well as in Aristotle and Aquinas. The classification is between presentations (*Vorstellung*), judgments (*Urteil*), and "relations of feeling" (*Gemütstätigkeit*) or "phenomena of interest" (*Interessephänomene*), as he also calls them, such as love and hate. These three types of mental acts have different functions in mental life. They are not of equal footing but depend on each other in specific ways. Brentano's fundamental thesis was that all mental acts and processes are either presentations or founded on presentations. Or, as Brentano puts it, "[i]t is impossible for conscious activity to refer in any way to something which is not presented" (Brentano, *ibid.*, 198).

[35] In a letter to Ehrenfels, Mach emphasizes his role in the development of this theory and points out that "he had already put forward the main ideas in an earlier paper – albeit in a more psychological way, in terms of a theory of 'muscular sensation'." The paper to which Mach refers was titled "Bemerkungen zur Lehre vom räumlichen Sehen" and it was published in 1865 (Kevin Mulligan and Barry Smith, "Mach and Ehrenfels: The Foundations of Gestalt Theory," in Barry Smith (ed.), *The Foundations of Gestalt Theory* [München, Wien: Philosophia Verlag, 1988], 125).

[36] Christian von Ehrenfels, "On 'Gestalt Qualities'," in Barry Smith (ed.), *Foundations of Gestalt Theory* (München, Wien: Philosophia Verlag, 1988), 93.

[37] Ehrenfels, *ibid.*, 83.

[38] Gurwitsch ignores Ehrenfels' explicit attempts to avoid the concepts of summation and argues that for Ehrenfels the Gestalt was a mere sum of its elements (Gurwitsch, *The Field*, 59, 84).

a Gestalt quality which binds these elements together. If we change the key and transpose the tones, Ehrenfels argued, following Mach, then all the elements are changed, but the melody is still recognized.[39] On the other hand, if we present the same notes but change their order, then the melody is lost.[40]

According to Ehrenfels, the origin of the Gestalt quality is not in intellectual activity nor in the imagination,[41] but in sense-perception or sensory presentation itself. Ehrenfels shared this view with Mach, even though he rejected Mach's attempt to explain the additional quality by reference to our kineasthetic sensations of our own sense-organs.[42]

Thus, both Ehrenfels and Mach argued against the traditional view according to which complexes are not given in sense-perception but generated or synthesized by some other mental faculty or activity, imaginative or intellectual.[43] In his *Contributions to the Analysis of the Sensations* (*Analyse der Empfindungen* 1886) – one of Ehrenfels' main sources – Mach stated:

> If two series of tones be begun at two different points on the scale, but be made to maintain throughout the same rations of vibration, *we recognize in both the same melody, by an act of sensation*, *just as directly* as we recognize in two geometrically similar figures, similarly situated, the same shape.[44]

Ehrenfels saw this as a plausible notion in Mach's discussion. He warned against taking the term "sensation" (*Empfindung*) with all its usual connotations,[45] but emphasized at the same time that Mach's proposal is true to experience:

> It becomes clear in the course of his discussion, however, that in using this perhaps not completely precise designation, Mach wished to give prominence to the immediacy of certain impressions and to *their independence from all intellectual processing on the part of the perceiving subject*.[46]

Both Gurwitsch and Merleau-Ponty celebrated this view as the important legacy of the early Gestaltists. In *The Field of Consciousness* (1964), Gurwitsch explains that Ehrenfels defended the idea that Gestalt forms pertain to the realm of sensibility, because Ehrenfels acknowledged with Mach that they are given to us with an immediacy comparable to that of simple percepts, such as colors and musical notes. Gestalts are not subsequent products of some higher mental functions, such

---

[39] Ehrenfels believed that this was Mach's understanding, but he was not completely sure of his interpretation (Ehrenfels, *ibid.*, 83–84).

[40] Ehrenfels, *ibid.*, 90.

[41] Ehrenfels, *ibid.*, 84.

[42] Ehrenfels, *ibid.*, 90, cf. Mulligan and Smith, "Mach and Ehrenfels," 126–28.

[43] Mach wrote: "In melodic as well as in harmonic combinations, notes whose rates of vibration bear to one another some simple ratio are distinguishable (1) by their agreeableness, and (2) by a sensation characteristic of this ratio," quoted in Ehrenfels, *ibid.*, 84; cf. Ernst Mach, *Beiträge zur Analyse der Empfindungen* (Jena, 1886), 130, *The Analysis of Sensations and the Relation of the Physical to the Psychical*, trans. C.M. Williams and S. Waterlow (New York: Dover, 1959), 287.

[44] Quoted in Barry Smith, "Gestalt Theory," 285.

[45] Ehrenfels, *ibid.* 82.

[46] Ehrenfels, *ibid.*, 83.

as imagination or intellect, but are given together with the other elements of sense perception.

Gurwitsch pointed out that all attempts to explain organization in experience by a special agency, principle, or activity of the mind face the problem of motivation: "Whatever the specific terms used in explanatory attempt, the question unavoidably arises as to the *clues* which might guide the organizing principle."[47] His main argument was that organization is an original, "autochthonous," feature of all contents of consciousness, perceptual and non-perceptual. Thus he proceeded to present his thesis of a universal formal pattern of organization, realized in all activities of consciousness regardless of their content.[48]

As pointed out above, Ehrenfels did not conceive of Gestalt as a whole that would include its constituents as dependent parts, but defined it as an additional, "new" quality of a complex of "mutually separable" presented elements. This was his explicit position and argument, even though he also claimed that the parts of the Gestalt are "fused into a whole",[49] which hinted at another direction.

Ehrenfels' concept of Gestalt was later problematized by Max Wertheimer, Wolfgang Köhler and Kurt Koffka, the later Gestalt theorists that founded the Berlin School at the beginning of the century.[50] In his *Gestalt Theory* (*Über Gestalttheorie* 1924), Max Wertheimer states:

> What is given me by the melody does not arise [. . .] as a secondary process from the sum of pieces as such. Instead, what takes place in each single part already depends upon what the whole is.[51]

So according to Wertheimer and his colleagues, Koffka and Köhler, the elements of the Gestalt form are determined and dependent on each other and on the overall structure. In Brentano's vocabulary, they are mutually non-separable parts or

---

[47] Gurwitsch, *The Field*, 54. In Gurwitsch's understanding the tacit acceptance of the constancy hypothesis had led earlier thinkers to look for the origin of organization outside perceptual consciousness, either in the physical reality or in some higher mental functions (*ibid.*, 52, cf. 70). Gurwitsch argues that all his predecessors, from Ehrenfels to Stump, Husserl, and the members of the Graz School, accepted, implicitly or explicitly, the constancy hypothesis. He states that it was not until the results of Wolfgang Köhler and Kurt Koffka in the 1910s, that the hypothesis was questioned (*ibid.*, 90–1, 161ff.).

[48] According to Gurwitsch, this pattern always involves three dimensions: a thematic center, a thematic field closely connected to the center, and a margin more loosely related to the center. Gurwitsch uses the concept "material relevance" to account for the relations between the different dimensions (Gurwitsch, *The Field*, 55–56). The similarity with Husserl's concept of horizon becomes obvious here.

[49] Ehrenfels, *ibid.*, 115.

[50] Wertheimer's early paper "Experimentelle Studien über das Sehen Bewegung" (1912) is considered the origin of the Berlin School and is often also presented as the starting shot of Gestalt-theoretical inquiries. Its topic was the so-called phi-phenomenon, a perceptual illusion in which two stationary but alternately flashing lights appear to be a single light moving from one location to the other.

[51] Max Wertheimer, "Über Gestalttheorie," *Philosophische Zeitschrift für Forschung und Aussprache* 1 (1925), 39–60; "Gestalt theory," trans. Willis D. Ellis, in *Source Book of Gestalt Psychology* (New York: Harcourt, Brace and Co, 1938).

moments of the whole. Thus, the constitutive elements do not sustain the decomposition of the form. When the structure is changed, then also its elements are lost.

In his introductory essay to *Foundations of Gestalt Theory* (1988), Barry Smith argues that we have here two different notions of Gestalt. For Ehrenfels and for his "more faithful" Austrian followers (in the Graz School), Gestalt is an additional quality that is associated with and belongs to the complex of "data given in experience." For the Berlin School, the collection of data "does not *have* a Gestalt, it *is* a Gestalt."[52] So on the one hand, the Gestalt is understood as a quality of "data complexes", and on the other hand, it is taken as an independent structure that makes possible such complexes and their constituent parts.

14.4 Merleau-Ponty's Critical Response to Gestalt Theories

In the introductory part of his *Phenomenology of Perception* (*Phénoménologie de la perception* 1945), Merleau-Ponty launches a Gestalt theorist argument against atomistic theories of sensation. More precisely, Merleau-Ponty refers to the Gestalt-psychological results of the Berlin School when arguing against the notion that the perceived thing is composed of simple and independent sensory units, i.e. sense-qualia.[53]

Merleau-Ponty finds this atomistic notion entertained both in the empiricist tradition of modern philosophy, represented by his contemporary positivists but originating in Hume and Locke, and in its intellectualist counter current, having its origin in Descartes' theory of perception. He even argues that the empiricist notion of sense-data is implicit in Descartes' rejection of sensations from the realm of clear and distinct ideas.[54] Thus, his argument is intended to be far-reaching and to lay fresh ground for our understanding of sense-perception.

After this constructive discussion and application of Gestalt-theoretical results, Merleau-Ponty states, however, that new ground can be broken only if Gestalt-theoretical results are reinterpreted within the methodological and conceptual framework of Husserlian phenomenology. This means that Gestalt theory has to be subjected to transcendental-phenomenological reduction, and all its ontological preconceptions – theoretical as well as practical, physicalistic as well as humanistic – have to be suspended.[55] Merleau-Ponty states this directly:

---

[52] Smith, "Gestalt Theory," 13; cf. Mulligan and Smith, "Mach and Ehrenfels," 131–32.

[53] Merleau-Ponty's references are to the figure-background phenomena, such as the Müller-Lyer arrows.

[54] For a more detailed discussion of Merleau-Ponty's reading of Descartes, see Sara Heinämaa, "The Living Body and its Position in Metaphysics: Merleau-Ponty's Dialogue with Descartes," in Dan Zahavi, Sara Heinämaa and Hans Ruin (eds.), *Metaphysics, Facticity, Interpretation* (Dordrecht, Boston, London: Kluwer, 2003), 23–48.

[55] In this respect Merleau-Ponty (as well as Gurwitsch) are closer to Ehrenfels than to the members of the Berlin School, because Ehrenfels already argued, with Brentano and against Mach, that

> But what Gestalt psychology lacks for the adequate expression of these perceptual rela-
> tionships is a set of new categories: it has admitted the principle [by rejecting the con-
> stancy hypothesis], and applied it to a few individual cases, but without realizing that a
> complete reform of understanding is called for if we are to translate phenomena accu-
> rately; and that to this end the objective thinking of classical logic and philosophy will
> have to be questioned, the categories of the world laid aside, the alleged self-evidence of
> realism placed in doubt, in the Cartesian sense, and a true "phenomenological reduction"
> undertaken.[56]

In a footnote he continues his critique by writing:

> Gestalt psychology cannot see that psychological atomism is only one particular case of a
> more general prejudice: the prejudice of a determinate being or of the world.[57]

Following Husserl, Merleau-Ponty argues that it is not enough to turn back to study
the perceptual experience and to describe the objectivities involved in it. Such a
return must be deepened by the phenomenological-transcendental reduction which
suspends, most importantly, the prejudice of the world.[58] The aim is not to dwell on
the givenness of the perceptual field, or the life-world, but to proceed to inquire into
the constitutive functions of consciousness.

> [...] the return to perceptual experience, in so far as it is a consequential and radical re-
> form, puts out of court all forms of realism, that is to say, all philosophies which leave
> consciousness and take as given one of its results.[59]

According to Merleau-Ponty, Gestalt psychology has not taken this step and has
not managed to break with naturalism, and thus suffers from naive notions of the
nature of being (the being of the world).[60] As a special science or theory of psy-
chic reality, Gestalt psychology may dispense with such radical inquiries. But if the
aim is to provide a foundation for the sciences and for the scientific enterprise as a
whole, as is claimed for example by Koffka in *Principle of Gestalt Psychology*, then
Gestalt psychologists must subject their explanations as well as their findings to a
radical philosophical investigation which concerns the meaning of being, its con-
stitution and its genesis. Only on this condition, Merleau-Ponty claims, can Gestalt

---

we should not allow physics direct our choice and development of the concepts that we use in
accounting for the processes and structures of the psychic (Ehrenfels, *ibid.*, 86–88).

[56] Maurice Merleau-Ponty, *Phénoménologie de la perception* (Paris: Gallimard, 1993 [1945]), 60;
Phenomenology of Perception, trans. Colin Smith (London and New York: Routledge, 1995), 49.

[57] Merleau-Ponty, *ibid.*, 62–3, n1; 51, n1. This is also a reply to Koffka who, in his *Principles of
Gestalt Psychology*, "tried to show that [Husserl's] attack on psychologism leaves Gestalt psychol-
ogy untouched" (Merleau-Ponty, *Phénoménologie*, 62; n1, *Phenomenology*, 50, n1).

[58] Several commentators have argued that Merleau-Ponty rejects either Husserl's transcendental-
phenomenological reduction or his eidetic reduction or both reductions. For a contrary argument,
see Sara Heinämaa, "From Decisions to Passions," in Ted Toadvine and Lester Embree (eds.),
Merleau-Ponty's Reading of Husserl (Dordrecht, Boston, London: Kluwer, 2002), 127–46.

[59] Merleau-Ponty, *Phénoménologie*, 58; *Phenomenology*, 47.

[60] For Merleau-Ponty's discussion and critique of the Berlin School, see his "Le métaphysique
dans l'homme," in *Sens et non-sens* (Paris: Gallimard, 1996 [1966]), 104–06; "The Metaphysical
in Man," in *Sense and Non-Sense*, trans. Hubert L. Dreyfus and Patricia Allen Dreyfus (Evanston,
Illinois: Northwestern University Press, 1964), 85–6.

theory claim to have epistemological or ontological relevance. The description of the phenomenal field must thus be deepened by an investigation of the transcendental realm:

> [...] we had to begin our examination of perception with psychological considerations. If we had not done so, we would not have understood the whole meaning of the transcendental problem [...] We had to frequent the phenomenal field and become acquainted, through psychological descriptions, with the subject of phenomena, if we were to avoid placing ourselves from the start, as does reflexive philosophy, in a transcendental dimension assumed to be eternally given, thus by-passing *the full problem of constitution*.[61] [...] But now that the phenomenal field has been sufficiently circumscribed, let us enter this ambiguous domain and let us steady our first steps with the psychologist, until the psychologist's self-scrutiny leads us, by way of a second-order reflection, to the phenomenon of the phenomenon, and decisively transforms the phenomenal field into a transcendental one.[62]

With these words, Merleau-Ponty ends the first, introductory section of his *Phenomenology* and begins his actual analyses. The first part of the book proper is titled "The Body" (*Le corps*), and what follows is an extensive discussion of the sensory, motor, affective, sexual, and expressive operations of the body as it is experienced and lived in its constant interaction with things and other living bodies.[63] Several interpreters – Gurwitsch among them – have argued that in these sections Merleau-Ponty rejects Husserl's transcendental aspirations, abandons the constitutive analyses of classical phenomenology, and offers instead a mundane theory of bodily subjectivity or motor agency.[64] As Merleau-Ponty's discourse revolves around the body, many commentators assume that he turns to, or falls back on, some form of mundane philosophy – realism, pragmatism, or even naturalism – for how could the body function as a constitutive ground in the framework of the philosophy of consciousness?[65] Contrary to this received view, the last part of the chapter shows that by thematizing the lived body Merleau-Ponty does not propose any new type of

---

[61] By "reflexive philosophy" and its "eternal dimension," Merleau-Ponty refers to Kant's transcendentalism and the Kantian methodology, which according to him cannot but bypass the constitutive of problems of personality, intersubjectivity, intercorporeality, and passivity (cf. *Phénoménologie.*, 75–6; *Phenomenology*, 61–3).

[62] Merleau-Ponty, *ibid.*, 76–7; 63, translation modified and emphasis added.

[63] For an introduction to Husserl's and Merleau-Ponty's phenomenologies of the body, see Sara Heinämaa, *Toward a Phenomenology of Sexual Difference: Husserl, Merleau-Ponty, Beauvoir* (Lanham, Boulder, New York, Oxford: Rowman & Littlefield, 2003).

[64] Aron Gurwitsch, "Review of *Phénoménologie de la perception*," *Philosophy and Phenomenological Research* 10 (1950), 444; *The Field*, 305; *Studies in Phenomenology and Psychology*, 349 n52; cf. Dillon, "Gestalt Theory," 457; Lester Embree, "Gurwitsch's Critique of Merleau-Ponty," *Journal of the British Society for Phenomenology* 12 (1981), 151–63; Ted Toadvine, "Phenomenological Method in Merleau-Ponty's Critique of Gurwitsch," *Husserl Studies* 17 (2001), 195–205; Beata Stawarska, "Anonymity and Sociality: The Convergence of Psychological and Philosophical Currents in Merleau-Ponty's Ontological Theory of Intersubjectivity," *CHIASMI International* 5 (2004), 295–309.

[65] Some commentators argue that instead of a transcendental egology, Merleau-Ponty proposes a transcendental theory of the flesh, see, e.g., Gary Madison, *The Phenomenology of Merleau-Ponty* (Athens: Ohio University Press, 1981 [1973]); Martin C. Dillon, *Merleau-Ponty's Ontology* (Bloomington and Indianapolis: Indiana University Press, 1988).

realism or naturalism but follows the guidelines that Husserl gave to transcendental phenomenological inquiries.

To begin to understand Merleau-Ponty's insistence on the radicalism[66] of Husserl's transcendental phenomenology and its constitutive analyses, it is important to notice a historical fact which concerns Merleau-Ponty's reading of Husserl's works. Whereas most of Merleau-Ponty's colleagues and contemporaries, Gurwitsch and Sartre included, knew only Husserl's early publications, primarily the *Logical Investigations* and the first volume of *Ideas*, Merleau-Ponty also studied Husserl's later works and his unpublished manuscripts.[67] In *Phenomenology of Perception*, Merleau-Ponty refers repeatedly to Husserl's three great works from the 1920s and 1930s, *Formal and Transcendental Logic* (*Formale und transzendentale Logik* 1929), *The Crisis of European Sciences* (*Die Krisis der europäischen Wissenschaften* 1936–1936, 1954), and *Experience and Judgment* (*Erfahrung und Urteil* 1939).

This historical fact is important both thematically and methodologically: in these works, Husserl tackled a set of new problems – passivity, embodiment, and intersubjectivity – but he also introduced a new of type of transcendental inquiry, so-called *genetic phenomenology*. This methodological shift does not mark a return to causal or empirical-historical explanations (unlike Brentano's "genetic psychology");[68] rather its aim is to disclose the temporal and motivational relations between intentional experiences. The task is to explicate the *necessary* temporal ordering of different intentional and lived experiences, the *essential* steps in the temporal layering of the sense of objectivity. In *Formal and Transcendental Logic*, Husserl explains his new approach as follows:

> It is the essential property of such products [experiential structures] that they are senses (*Sinne*) which bear a sort of historicity in themselves as the sense implications of their genesis; that in them a sense refers back by degrees to an original sense and to the noematic intentionality belonging to it; and that we can therefore investigate every sense formation for the *sense history essential* to it.[69]

---

[66] In addition to the quotations given above, see, e.g., Merleau-Ponty, *Phénoménologie.*, 253, 278–89, 418–19; *Phenomenology*, 219, 240–42, 364–65.

[67] An illuminating overview of Merleau-Ponty's reading of Husserl is given by Ted Toadvine, "Merleau-Ponty's Reading of Husserl," in Ted Toadvine and Lester Embree (eds.), *Merleau-Ponty's Reading of Husserl* (Dordrecht, Boston, London: Kluwer, 2002), 227–86.

[68] Husserlian genealogy should be kept separate from genealogy in Nietzsche's sense. Despite their shared interest in the plurality of experience, and their common fight against false universalism and existential prejudices, the two genealogies remain very different insofar as Husserl's approach is *eidetic* and Nietzsche's approach is explicitly focused on historical and cultural *contingences*. On Nietzschean genealogy, see Michel Foucault's clarifying essay, "Nietzsche, la généalogie, l'histoire," in *Dits et écrits 1954–1988, I: 1970–1975* (Paris: Gallimard, 1998), 136–56; "Nietzsche, Genealogy, History," trans. J. Harari, in Paul Rabinow (ed.),: *The Foucault Reader* (Harmondsworth: Penguin, 1984), 76–100.

[69] Husserl, *Formale*, 184; cf. *Erfahrung und Urteil: Untersuchungen zur Genealogie der Logik*, ed. Ludwig Landgrebe (Hamburg: Felix Meiner Verlag, 1999 [1939]), 19–20, 45ff.; *Experience and Judgment: Investigations in the Genealogy of Logic*, trans. James S. Churchill and Karl Ameriks (Evanston: Northwestern University Press, 1973), 26, 47ff.; *Die Krisis der europäischen*

In genealogical investigations, the main interest of the phenomenologist is no longer in the stable, fixed, or "ready made" structures of experience, but is instead in the dynamically evolving and developing senses (*Sinne*) which pile up one upon the other within the inner time of consciousness.[70] Already established layers of sense are enriched and enlarged by new sense components. In *Cartesian Meditations*, Husserl gives a concrete example by describing the institution of tool experience: "The child who already sees physical things understands, let us say, for the first time the final sense of scissors; and from now on he sees scissors at the first glance *as* scissors."[71] The sense 'material thing' is thus enriched by the senses 'utensil' and 'instrument'. Husserl calls such a process of gradual temporal establishment and enrichment of sense "sedimentation". This idea is the key to the understanding of Merleau-Ponty's discourse of the body and its perceptual environment.

In this respect it is important to pay attention to Merleau-Ponty's repeated references to Husserl's manuscript of the second volume of *Ideas*, entitled *Studies in the Phenomenology of Constitution*. This work is not an introduction or a presentation of the phenomenological method but offers detailed constitutive analyses of different types of objectivities, from mere physical things to living bodies (animal and human), cultural objects and artifacts (utensils, instruments, and works of art), and expressive means (verbal and gestural). Husserl's manuscript gave Merleau-Ponty a rich stock of models for his own constitutive inquiries. Moreover, Husserl developed a phenomenological account of constitutive ordering. He argued that the experience of one's own living, sensing, and moving body provides the basis for all experiences of worldly objects. The constitution of perceivable things depends on the transcendental fact that one's own body can appear both as sensing (subject) and as sensed (object) at the same time.[72] The experience of one's own sensing-sensed body provides the foundation for the constitution of the original sense of 'living,' and this sense in turn gives the foundation for the meaning of an 'alien' living body. Together, living bodies – animal bodies and human bodies – take part in the constitution of the perceivable world.[73] Thus, the perceivable world is not a creation of a

---

Wissenschaften und die transzendentale Phänomenologie, 48–52, 123ff.; *The Crisis of European Sciences*, 48–52, 121ff.

[70] For Husserl's genetic account of the personal self, see Tetsuya Sakakibara, "Das Problem des Ich und der Ursprung der genetischen Phänomenologie bei Husserl," *Husserl Studien*, 14 (1997), 2–39; Sara Heinämaa, "Selfhood, Consciousness, and Embodiment: A Husserlian Approach," in Sara Heinämaa, Pauliina Remes and Vili Lähteenmäki (eds.), *Consciousness: From Perception to Reflection in the History of Philosophy* (Dordrecht: Springer, 2007), 311–28.

[71] Husserl, *Cartesianische Meditationen und Pariser Vorträge*, 141; *Cartesian Meditations*, 111.

[72] Husserl, *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Zweites Buch*; 147ff.; *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, Second Book*; 155ff.; cf. Merleau-Ponty, *Phénoménologie*, 109ff., 364–365, 406–07; *Phenomenology*, 93ff., 315–16, 353–54.

[73] This is the main argument of the fifth Cartesian Meditation. See also *The Crisis of European Sciences*, where Husserl writes: "What properly and essentially makes up the character of a living body I experience only in my own living body, namely, in my constant and immediate holding-sway [over my surroundings] through this body alone. Only it [my own body] is given to me originally and meaningfully as 'organ' and as articulated into particular organs [. . .] Obviously it is only in

solitary, bodiless subject but an intersubjective and intercorporeal accomplishment. The objective realms of the natural and human sciences are established on the basis of the shared world of perception.

Merleau-Ponty used the central concepts of Husserl's genetic phenomenology, most importantly the concept of *sedimentation* (*Sedimentierung*, *sédimenation*), to account for the necessary temporal relations between the different senses of perceivable objects: one's own living body, the body of the other, the natural thing, and the cultural artifact. First, he argued that the *body schema* (*scheme corporel*) – the kinaesthetic and synaesthetic form of the living body – is the source of meaning and a necessary condition for all the objectivities that we find in the world.[74] Even the senses of geometrical objects and those of linguistic entities can be tracked down to this source through the constitution of the human body: "Our body, to the extent that it moves itself about, that is, to the extent that it is inseparable from a view of the world and is that view itself brought to existence, is the condition of possibility, not only of the geometrical synthesis, but of all expressive operations and all acquired views which constitute the cultural world."[75]

But more ambitiously Merleau-Ponty also wanted to account for the constitutive origin of the basic form of the human body. His solution is an application of Husserl's idea of sedimentation but it also extends this idea over new dimensions. While arguing that the different senses of objectivity are layered one upon the other, Merleau-Ponty suggests that also our body schema is a sediment. It is not, however, a sediment of our own sense-giving activities, it does not result from our personal operations or actions. Rather we have received or inherited this primordial form from our intentional ancestors, that is, from other living subjects who precede us in time. This means that our perceptions are not original accomplishments but reactivations or re-constructions; they depend on the constitutive work already done by others. Thus Merleau-Ponty argues that phenomenological reflection, when performed in a radical way, discloses our perceptions as moments in an entire *tradition* of sensing and perceiving. We find this idea explained in all the main sections of Merleau-Ponty's *Phenomenology*. In the chapter on *Sense Experience* (*Le sentir*), we read:

---

this way [i.e. by having sense organs] that I have perceptions and, beyond this, other experiences of objects in the world. [. . .] Only through my originally experienced holding-sway, which is the sole original experience of living-bodiliness as such, can I understand another body as a living body in which another 'I' is embodied and holds sway [. . .] Only in this way do other ego-subjects firmly belong to 'their' bodies for me and are localized here or there in space-time" (*Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie*, 220–21; *The Crisis of European Sciences*, 217–18).

[74] For an introduction to the theory of the body schema, see Shaun Gallagher and Jonathan Cole, "Body Image and Body Schema," in Donn Welton (ed.), *Body and Flesh: A Philosophical Reader* (Oxford: Blackwell, 1998).

[75] Merleau-Ponty, *Phénoménologie*, 455; *Phenomenology*, 388.

> My act of perception, in its naiveté, does not itself bring about this [aesthetic] synthesis; it takes advantage of work already done, of a general synthesis constituted once and for all, and this is what I mean when I say that I perceive with my body or my senses, since my body and my senses are precisely that familiarity with the world born of habit, that implicit or sedimentary body of knowledge. [...] The *person who* perceives is not spread out before himself as a consciousness must be; he has a historical density, he takes up a perceptual tradition and is faced with a present.[76]
>
> Between my sensation and myself there is always the thickness of some *original acquisition* which prevents my experience from being clear to itself. I experience the sensation as a modality of a general existence, one already devoted to a physical world and which runs or streams through me without my being its originator.[77]

Merleau-Ponty does not propose that the human body is an independent objectivity – independent of the constitutive activities of intentional subjects (his philosophy is not realistic). Nor does he suggest that the human body is the ultimate agent of all constitution. Instead he argues that the form or the schema of my living body has originated with earlier, alien forms of conscious life. Thus, our bodies are not constituted by our personal activities, decisions, volitions, or judgments; but this transcendental fact does not mark the limit of the constitutive process. The form of our body is received from earlier phases of intentional history; it is pre-personal and pre-historical.

This account is certainly controversial from the point of view of classical phenomenology. But it is important to realize that its unorthodox character does not come from an abandonment of the concept of constitution. On the contrary, Merleau-Ponty moves to the limits of classical phenomenology insofar as he extends or stretches the idea of constitution from its primary coverage over one consciousness, and the community of synchronic subjects, to a whole history of diachronic consciousnesses. In other words, Merleau-Ponty's transcendental phenomenology proceeds from *genetic* investigations to *generative* inquiries.[78] It proposes that objectivity is not established merely in the interaction between several synchronous consciousnesses but depends on generations of diachronic others. This obviously raises the question: How is Merleau-Ponty able to account for trans-generational (and even trans-human) temporal references without making compromises over Husserlian methods? *Phenomenology of Perception* does not answer this question; but it formulates and motivates the question with exceptional vigor and thus urges us to search for answers.[79] At the end of the chapter on *Temporality*, Merleau-Ponty writes:

---

[76] Merleau-Ponty, *ibid.*, 275; 238. Cf. Sara Heinämaa, "Personality, Anonymity and Sexual Difference," in Helen Fielding, Dorothea Olkowski and Christina Schües (eds.), *Time in Feminist Philosophy* (Indiana University Press, forthcoming in 2009).

[77] Merleau-Ponty, *ibid.*, 250; 216.

[78] For an extensive account of the concepts 'genetic' and 'generative', see Anthony Steinbock's *Home and Beyond: Generative Phenomenology After Husserl* (Evanston, Illinois: Northwestern University Press, 1973).

[79] Merleau-Ponty, *ibid.*, 417–18; 363–64.

As my living present opens upon a past which I nevertheless am no longer living though, and on a future which I do not yet live, and perhaps never shall, it can also open on to temporalities outside of my living experience and acquire a social horizon, with the result that my world is expanded to the dimension of that collective history which my private existence takes up and carries forward.[80]

---

[80] Merleau-Ponty, *ibid.*, 495; 433.

Chapter 15
Philosophy of Mind with and Against Wittgenstein

Frederick Stoutland

Abstract The chapter investigates the discussion of twentieth century analytical philosophy of mind, focusing on Wittgenstein's and Davidson's criticism of physicalism. Wittgenstein's most significant contribution was his rejection of the Cartesian mind–body distinction, the accompanying conception of philosophy, and the idea of primacy of action over intellection. Physicalism, the view that the mental must either be an explanatory physicalist theory or vindicated by such a theory, is exemplified in the chapter by the theories of Fodor and Churchland. Fodor accepts the mental because it can be made scientifically respectable; Churchland rejects it because it cannot. The chapter argues that both views assume the Cartesian distinction (manifested in what they call "folk psychology") between the physical – conceived as physicalistic – and the mental – conceived as a realm of inner entities. Davidson, however, rejects physicalism and the Cartesian mental–physical distinction that underlies it. By elucidating his philosophy of action the chapter shows that it is unjustified to equate his monism with physicalism. Davidson makes a sharp distinction between rational explanations of action (which are interpretative, normative, and first-person) and physical scientific explanations. Most importantly, he does not claim that rational explanations of action must be grounded in physical laws. In spite of Davidson's recurrent reference to theories, his work is non-theoretical in Wittgenstein's sense.

Keywords Mind–body relation · physicalism · philosophy of action · rule-following · Davidson · Wittgenstein

This chapter presents a view of what in twentieth-century philosophy of mind done in the analytical tradition is likely to be taken seriously by philosophers in, say, the twenty-third century. I discuss what I take to be high points in twentieth-century philosophy of mind, three in particular, namely Wittgenstein, physicalism, and Davidson. I consider Wittgenstein the preeminent philosopher of the century, whose

F. Stoutland
St. Olaf College (Minnesota) and University of Uppsala

work will be more and more appreciated and assimilated as time goes by. I consider physicalist philosophy of mind to be a noteworthy continuation of the Hobbesian tradition that will always be with us, but I doubt that any of its individual practitioners will stand out as more historically significant than such nineteenth-century physicalists as Jean Cabanis or Ludwig Buechner (whose 1855 work, *Force and Matter*, went through 16 editions). I consider Davidson a splendid example of how to do philosophy of mind in a manner inspired by Wittgenstein without being a Wittgensteinian, someone history may remember as the most important analytical philosopher of mind.

Before continuing, I want to say a word about analytical philosophy itself, whose boundaries are becoming less and less clear. Wittgenstein differed from paradigm analytical philosophers in such fundamental ways that historians of philosophy will probably cease to regard him as an analytical philosopher. Nevertheless, he surely belongs to the analytical *tradition* because of his philosophical education (Frege and Russell), his influence on analytical philosophers (however much it rested on misunderstanding), and the fact that he worked in Cambridge when it was the center of analytical philosophy. Ever since logical positivism ceased to be influential, physicalists have constructed the same kind of metaphysical theories philosophers like Moore thought it was the task of analytical philosophy to stamp out; what makes such metaphysics analytical is that it is done by philosophers who were trained in the analytical tradition. Davidson rejected such metaphysical theorizing, his manner of working being more analytical in the classical sense than either Wittgenstein's or the physicalists'; his views, however, are closer to Wittgenstein's than to those held by most analytical philosophers.

15.1 Wittgenstein and Philosophy of Mind

What themes in Wittgenstein's work in philosophy of mind will be significant in the long term? There are, I suggest, five, which are interconnected in intricate ways.

1. His conception of a philosophical investigation.
2. His rejection of the Cartesian conception of the distinction between the mental and the physical.
3. His reworking of Kant's notion of the inseparability of intuitions and concepts.
4. His conception of action as prior to intellection.
5. His renewal of the notion that the world itself is both evaluatively and motivationally significant.

(1) The best account of Wittgenstein's conception of philosophy is his own in *Philosophical Investigations*, especially paragraph 109:

> It was true to say that our considerations could not be scientific ones. It was not of any possible interest to us to find out empirically "that, contrary to our preconceived ideas, it is possible to think such-and-such" – whatever that may mean [...] And we may not advance any kind of theory. There must not be anything hypothetical in our consideration. We must do away with all *explanation*, and description alone must take its place. And this

description gets its light, that is to say its purpose, from the philosophical problems. They are, of course, not empirical problems; they are solved, rather, by looking into the workings of our language, and that in such a way as to make us recognize those workings: *in spite of* an urge to misunderstand them. The problems are solved, not by reporting new experience, but by arranging what we have always known.[1]

This should be supplemented with paragraph 122, which reads (in part):

A main source of our failure to understand is that we do not command a *clear view* of the use of our words. – Our grammar is lacking in this sort of perspicuity. A perspicuous representation produces just that understanding which consists in 'seeing connections'. [. . .] The concept of a perspicuous representation is of fundamental significance for us. It earmarks the form of account we give, the way we look at things.[2]

Wittgenstein's conception of philosophy is crucial for understanding his work, and it will be central to any historical account of his philosophical achievement. It also raises more hostility than any other facet of his work, and I want to mention two reasons for this. One is that his conception of philosophy cuts against the grain of the predominance in our culture of a scientistic mentality that shows up in the fact, as he himself put it in the *Blue Book*, that "Philosophers constantly see the methods of science before their eyes, and are irresistibly tempted to answer questions in the way science does."[3] I'll give an example of how this works when I come to physicalism. The other is that his view of philosophy is regarded as a "quietism" that aims at – even announces – the "end of philosophy", which I think seriously misunderstands his view. Here is a more accurate characterization of it from Marie McGinn.

The things which we are doomed to misunderstand when we take up a theoretical attitude toward them are, then, just those things "that we know when no one asks, but no longer know when we are supposed to give an account of (them)" (Augustine). [. . .] What we are concerned with when we ask questions of the form 'What is time?', 'What is meaning?', 'What is thought?' is the nature of the phenomena which constitute our world [. . .] and in asking these questions we express a desire to understand them more clearly. Yet in the very act of framing these questions, we are tempted to adopt an attitude toward these phenomena which, Wittgenstein believes, makes us approach them in the wrong way, in a way which assumes that we have to uncover or explain something. [. . .] As soon as we try to catch hold of them in the way our questions seem to require, we find that we cannot do it; we find that we "no longer know" [. . .] We think that the fault lies in our explanations and that we need to construct ever more subtle and surprising accounts. [. . .] The real fault [. . .] is not in our explanations, but in the very idea that the puzzlement we feel can be removed by means of a discovery. What we really need is to turn our whole inquiry round and concern ourselves, not with explanation or theory construction, but with description. [. . .] For our puzzlement concerned the nature or essence of particular phenomena [. . .] and this puzzlement is removed. "not by giving new information, but by arranging what we have always known". [. . .] [Our aim should be] a kind of understanding which consists in seeing

---

[1] Ludwig Wittgenstein, *Philosophical Investigations* (Oxford: Basil Blackwell, 1958), 47.

[2] *Ibid.*, 49.

[3] Ludwig Wittgenstein, *The Blue and Brown Books* (Oxford: Basil Blackwell, 1960), 18.

a pattern or form in what is there before our eyes, but which we had previously neglected or overlooked. [...] Everything we need to understand is already there and only needs to be arranged correctly.[4]

Let me make two points against the background of this quotation. The first concerns what Wittgenstein *means* in saying that philosophers ought not to construct theories. He does not mean that we should not give arguments or come to conclusions. The constructive task of philosophy is to arrange what we have always known in a perspicuous representation, and that is arduous work and real philosophy. What we have always known is not obvious, and the most significant matters are embedded not in what we believe but in everyday practice and hence have to be made explicit. This requires distinguishing what we have always known from what we *think* we have always known, it requires *articulating* what we have always known, and it requires arranging it *properly*. The latter may require re-arrangement, which must be defended by argument and by showing *how* it resolves the philosophical problems that give purpose to the inquiry.

The second point concerns Wittgenstein's objections to theory construction, two in particular. One is that philosophers who construct theories invariably play fast and loose with the phenomena they try to explain. They think it is *obvious* what needs to be theoretically explained and hence are content to begin with platitudes, intuitive judgments, folk theories, and the like. They do not take the time to reflect on the phenomena themselves and the role they play in our thought, discourse, and action, and hence what they seek to explain is often the construct of some theory whose credentials are never examined. The other is that philosophical theories simply fail to explain what they claim to explain, the basic reason being that there is nothing to explain in the quasi-scientific sense of explanation that is their model. As philosophers, we must explain in the sense of *elucidate* but – and this is his distinctive claim – if we elucidate free from theoretical preconceptions about what there *must* be and how it *must* work, we will come to see that what theories purport to explain do not stand in need of such explanation.[5]

(2) Wittgenstein's most important substantive contribution to philosophy of mind is his rejection of the Cartesian distinction between the physical and the mental. This distinction conceives of the *physical* as consisting only of what plays an essential role in the new physics, a physics purified of the teleological, intentional, and normative terms of Aristotelean physics. The physical consists, that is, only of what can be specified in *physicalistic* terms that have no motivational or evaluative significance, knowledge of which must be based on observation any person is capable of and which is not interpretive but cumulative – claims can be settled and then added to. The *mental* is everything that takes place within ourselves so that we are

---

[4] Marie McGinn, *Wittgenstein and the Philosophical Investigations* (London: Routledge, 1997), 18, 26. This book is an extraordinarily illuminating discussion of central themes in the *Investigations*.

[5] For further discussion of this matter, see my "Analytical Philosophy and Metaphysics" in Sami Pihlström (ed.), *Wittgenstein and the Method of Philosophy* (Helsinki: *Acta Philosophica Fennica*, vol. 80, 2006).

immediately aware of it *qua* object of consciousness. It comprises internal states that can be known immediately to (and only to) their possessors through introspection, states which are so independent of behavior that they can be ascribed to immaterial souls (or brains in a vat).

It should be noted that Descartes himself had a third category, namely the mind–body union, which has features that belong neither entirely to the physical nor entirely to the mental, the latter being a distinction reached by *abstraction* from the mind–body union. What I call the Cartesian distinction does not, therefore, work in the same way for Descartes as for the Cartesians. It is the latter that Wittgenstein criticized, a point to which I shall return.

Wittgenstein's objection to the Cartesians was not that they made a distinction between the physical and the mental. He was not a behaviorist but simply took for granted a distinction between behavior and mental states – between pain behavior and pain, for example. He agreed with the Cartesians that there is an irreducible asymmetry between the ascription of psychological predicates to others and to oneself: I can doubt whether you are in pain (or are just pretending) but cannot doubt whether I am in pain, I can try to find out whether you are really in pain but not whether I am, I can express my pain but cannot express yours, and so forth.

His objection was two-fold. First, that Cartesianism misconstrued the distinction between the physical and the mental as a distinction between physicalistically described behavior whose features are observable to anyone with working sense organs, and internal states whose features are directly knowable only to their possessors and only by introspection. On this construal, a description of someone as suffering and in agony from pain must be broken into two parts: a description of purely physical happenings observable to anyone, and a description of internal happenings known only introspectively. Wittgenstein thought this was a distorted and obscuring way to describe either behavior or mental states, a form of description we give only because we are captives of a Cartesian theory about how things *must* be. He held that pain behavior is not even observable as such to creatures who have no concept of pain, no sense of what it is to suffer, no instincts to comfort the one afflicted. And he held that pain is not something we recognize in ourselves simply by bare introspection: recognition of it as pain requires that we are competent participants in the language in which pain plays its characteristic role.

Second, he objected to Cartesian attempts to explain *why* we distinguish the physical and the mental as we do – attempts not to *elucidate* the distinction but to *explain* it. Cartesians attempt to explain, for instance, the asymmetry in our ascription of mental states to ourselves and to others by appealing to the doctrine that mental states are internal and known only by introspection. That is supposed to *explain* why I can *doubt* that you are in pain and not doubt that I am, or why I can *express* my pain but not yours. Wittgenstein undertakes to show, in ways too intricate to spell out here, that the explanations are empty: if we work them through in particular cases, they fall apart and are seen, at best, simply to redescribe the phenomena, at worst, to be incoherent.

The failure of such explanation should not disappoint us but help us to see that there is nothing here that needs to be explained in that way. As Marie McGinn puts

it: "This complexity in the grammar of our sensation concepts is not something that needs to be explained [. . .] but is something which *in itself* reveals the fundamental distinction between sensations and behavior."[6] Philosophical reflection on our sensation concepts no doubt raises problems we find inescapable and difficult, but resolving *those* problems does not call for explanatory theory but for perspicuous representation.

Let me consider one more example of Wittgenstein's objection to the Cartesian distinction, namely, the so-called "rule-following considerations". Here the distinctive feature of our practice that the Cartesian aims to explain is that by understanding a word (grasping its meaning), we are able to use it in diverse contexts over time, or (another example of the same phenomenon) by grasping a rule, we are able to continue a series indefinitely or compute any number of sums. Wittgenstein has no objection to what I just said. In understanding the word "cube", for example, we are able to use it to refer to all sorts of objects that really are cubes, and by understanding the word "red" we can pick out objects which are red. By grasping the add-two rule, we are able to make the correct move from any natural number to its second successor, and by knowing how to add, we can compute, for instance, the sum of $68 + 57$ and come up with *the* correct answer, namely 125. Wittgenstein was not a skeptic about the meaning of words like "cube", "red", "add", or "plus" or about our capacity to determine the shapes or colors of objects or the truths of arithmetic. What he was a skeptic about was Cartesian attempts to *explain* why our understanding of words and rules determines the ways we use words and apply rules.

These attempts at explanation take at least three forms. One regards understanding as an inner state of grasping a meaning or rule that is at the same time an immediate grasp of all that is required by the meaning or the rule. Another takes understanding to be a mental mechanism that guides us whenever we use a word or apply a rule. A third thinks of understanding in terms of a Platonic vision of logical rails laid out before us: to understand a meaning or rule is to be logically compelled to use a word or follow a rule in the correct way. Wittgenstein argued, again in ways too intricate for discussion here, that if we consider exactly what these attempts at explanation claim about any particular case, we will see that they are empty. Understanding considered as an inner state, a mental mechanism, or a set of logical rails cannot guide us as it is said to do, for we equally need guidance in being guided. This is an intuitive way of putting Kripke's point about "plus" and "quus": the grasp we have of "plus" may require us to go on doing sums *in the same way* but it need not require us to understand "doing sums in the same way" *in the same way*.[7]

Kripke gave a nice exposition of Wittgenstein's critique of Cartesian explanations of how understanding determines use but in so doing he missed Wittgenstein's main point. He took Wittgenstein's critique to be a skeptical argument that there is no such thing as a correct use of a word or a correct application of a rule, so that what the Cartesian attempts to explain is an illusion. On Kripke's reading, there is no

---

[6] McGinn, *Wittgenstein and the Philosophical Investigations*, 123.

[7] Cf. Saul Kripke, *Wittgenstein on Rules and Private Language* (Oxford: Basil Blackwell, 1982).

such thing as understanding the rules of arithmetic and thereby coming up with 125 as the correct answer to 68 + 57. All we can say is that in *our community* we *call* 125 the correct answer, for, he argues, Wittgenstein has shown that the idea that 125 *is* the correct answer is empty.

But this is to turn Wittgenstein's skepticism about Cartesian *explanations* of how understanding a word or a rule enables us to apply it correctly into skepticism about whether there is any such thing as correct application, a skepticism Wittgenstein is said to counter by redescribing the phenomena in terms of a new theory, one which holds that assertions about correct application have communal assertability conditions but no *truth* conditions. On this reading, Wittgenstein's reflections on meaning and rule-following begin with a skepticism about their very possibility that is motivated by theoretical considerations and then tries to show that our ordinary practice can be sustained *in spite of* the truth of skepticism, provided we accept an explanation of the practice in terms of a verificationist theory of meaning.

For Wittgenstein, however, our ordinary practice, mathematical or otherwise, needs neither defense nor theoretical explanation, something that would be clear if we had a perspicuous representation of it. We make mistakes within it, but they are correctable from within it. What counts is not what processes go on in our mind or what formulae we grasp as we continue a series or do sums, for nothing can *compel* us to do one thing rather than another. That sense of compulsion is a myth, and hence its rejection does not mean that our application of words or rules is unconstrained by criteria of correctness (which would mean that there is no distinction between correct and incorrect). What counts is that we use words to communicate successfully, and apply rules so as to come up with correct answers, not by sheer luck or accident, of course, but by participation in the practice into which we were inducted as we learned to speak, to count, to add and subtract, to act for reasons, to contradict our peers, to coordinate our actions, and so on. The practice does not *make* what we do right nor does our competent participation in it *guarantee* that we will get it right; we nevertheless have the ability to get it right and to know when it is right.

Philosophical problems about our practice will arise as we reflect on it, but they will not be resolved by Cartesian theories about how inner states of understanding or mental mechanisms guide us in our speaking and calculating – nor by theories of any kind. Everything that we need to resolve the problems is already there in front of our eyes in what we have always known but have not been able to arrange perspicuously.

(3) Wittgenstein's conception of philosophy and his rejection of the Cartesian distinction between the physical and the mental are the most significant themes in his philosophy of mind. It is clear that they are not separable: his conception of philosophical inquiry supports his critique of the Cartesian view, and his rejection of that view shows why it is that our practice can be taken for what it is without a theoretical explanation of how it works. The other themes I mentioned can be seen as corollaries of these two and hence I will be brief.

The third theme is Wittgenstein's reworking of the Kantian idea that concepts without intuitions are empty and intuitions without concepts are blind. This is a

theme about intentionality – about what it is for us not only to *think* or *talk* about anything, rightly or wrongly, but to *respond* to anything as such and such, whether rightly or wrongly. What is decisive is to see that Wittgenstein did not mean it as a philosophical theory that, contrary to what one might expect, explains how we can think, talk, or respond to the world, other selves, and our own self.[8]

This theme shows up clearly in the so-called private language argument, whose aim is to show that introspection plays no role in defining psychological concepts because the notion of a private ostensive definition is simply an illusion. What a sensation or a feeling is, is to be understood not by looking within ourselves but by examining the roles the phenomena and our concepts of it play in our practice, which is conceptually structured. What this shows is that concepts of sensations or feelings are complex and are intricately related to our behavior – to the behavior that expresses them, that enables us to recognize them in others, that shows our concern, and so on. The point, again, is not to *explain* this complexity – to offer, for example, a theory about why there must be behavioral criteria for sensations or why we must be able to introspect them. It is this very complexity that shows what sensations and feelings are and how they are to be distinguished from behavior. Sensations may be *there* without concepts but without concepts they show us nothing about what they are (we can not recognize them) or what anything else is.

(4) The fourth theme concerns the nature of action and its status as prior to intellection. This is one dimension of his rejection of the Cartesian construal of the mental-physical distinction, a rejection that entails that action is not to be understood as the physical result of internal states or mechanisms, a point that lies back of his critique of understanding as an inner state or mechanism that guides our use of words or our procedures of calculation. The latter cannot be described in physicalistic terms since the notion of using a word or applying a rule has built into it concepts that already belong at the level of understanding and meaning. The same is true of action generally: the behavior that reasons can explain must be already conceptualized in non-physicalistic terms.

This suggests an equally important point about the *status* of action: since it cannot be understood as the physical result of internal states, it is more fundamental than intellection. What is fundamental is not what we think or believe but what we do – the practices in which we participate. This is not unique to Wittgenstein in twentieth-century philosophy: it is a motif in philosophers like Heidegger, Merleau-Ponty and John Dewey. Its articulation is, in my view, the most distinctive feature of twentieth-century philosophy.

(5) The last theme is Wittgenstein's renewal of the idea that the world is both evaluatively and motivationally significant. This is also a motif in the work of Heidegger,

---

[8] John McDowell argues that Kant didn't mean it that way either, that "no one has come closer than Kant to showing us how to find intentionality unproblematic" (except Wittgenstein). See his "Having the World in View," *The Journal of Philosophy* 95 (1998), 431.

Merleau-Ponty, and Dewey, for it is closely connected with the theme of action as more fundamental than intellection. The rejection of Aristotelean science meant (as Weber put it) the disenchantment of the world, a disenchantment that some philosophers, Descartes and Hume for instance, urge us to live with, and that others – Spinoza and Leibniz are examples – try to overcome by metaphysical theorizing. It is a theme in the background of much analytical philosophy of mind, which has taken for granted, as John McDowell puts it, that "Reason does not find meaning or intelligible order in the world: rather, whatever intelligible order there is in our world-picture is a product of the operations of mind, and those operations are themselves just some of what goes on in nature, in itself meaninglessly, as it were".[9]

Wittgenstein neither acquiesces in the disenchantment of the world nor gives us a theory explaining why it is evaluatively and motivationally significant. On his view, what disenchanted the world was not science as such but the position given to science by the Cartesian doctrine that made it the sole arbiter of the external world, with the rest going inside. This intellectualized language and action and robbed the world of its own evaluative and motivational significance. Both value and motivation were regarded as subjective – a result of our desires or other mental states being projected onto the world – not something there is in the world to be discovered. The reason for this is that these things are not manifest to mere physical observation or mental introspection. To understand how they manifest themselves, we need to restore the Kantian theme of the inseparability of concepts and intuitions, not as theory, but as what we have always known, perspicuously arranged.

15.2 Physicalism in the Philosophy of Mind

By "physicalism" I mean what Fred Dretske calls a "naturalistic theory of the mind", whose aim Jerry Fodor characterizes as finding a place for the mind "in a physicalistic view of the world." But what is a "physicalistic view of the world"? It can be taken weakly to mean two things: first, that we ought to believe what is entailed by the truths of natural science (physics in particular) and disbelieve what is inconsistent with them and, second, that if you destroy the physical, you destroy everything or, a little less dramatically, any change in the mental entails some change or other in the physical. But that *minimal* physicalism is something just about any philosopher (apart from supernaturalists) would endorse. My topic is *robust* physicalism, according to which an adequate account of the mental must either be, or be vindicated by, an explanatory theory from the natural sciences. On this view, there is a place for the mental in a physicalistic view of the world only if our concepts of the mind and the mental can be shown to function, or made to function, like explanatory concepts of a natural science (physics in particular).

---

[9] John McDowell, "Two Sorts of naturalism," in John McDowell (ed.), *Mind, Value, and Reality* (Cambridge: Harvard University Press, 1998), 174.

The coming and going of the diverse accounts of the mental that robust physical-
ists have proposed constitute a central part of the history of analytical philosophy of
mind. The most obvious motivation for these accounts is the unity of science project:
to show that any adequate explanation, whether in physics, biology, sociology,
history, or everyday life, shares a common form. Carnap's analytical behaviorism
was the first memorable such account in analytical philosophy of mind, one based
on the *positivist* construal of the unity of science, which assumed a sharp distinc-
tion between science as establishing truth and philosophy as analyzing meaning.
The unity of science project changed substantially with Quine's attack on the two
dogmas of empiricism, which undermined the positivist distinction between science
and philosophy and hence Carnap's behaviorist version of physicalism, which was
based on an analysis of the meaning of mental statements. The result was that the
positivist conception of the unity of science was replaced by a *naturalist* conception
that, on the one hand, regarded philosophy and science as continuous and, on the
other, regarded science and hence philosophy as essentially theory construction.
This paved the way for the robust physicalism that now dominates analytical philos-
ophy – so-called "analytical metaphysics" whose task is to construct theories like
those in the natural sciences and defend them by showing that they are better than
alternative theories in explaining the data and in resolving difficulties raised by the
theories themselves.[10]

The two best known physicalists of this kind are Paul Churchland and Jerry
Fodor, who stand at opposite ends of the physicalist spectrum in that they disagree
on the scientific worth of what Churchland calls "folk psychology" and Fodor "com-
monsense psychology". Churchland thinks that folk psychology is a hopelessly de-
fective theory of our internal processes that is as outmoded as the Ptolemaic theory
of the universe or the phlogiston theory of heat.

> [...] folk psychology is a radically inadequate account of our internal activities [...] that
> will simply be replaced by a better theory of those activities. [...] It provides a positively
> misleading sketch of our internal kinematics and dynamics. [...] It suffers explanatory
> failures on an epic scale, it has been stagnant for at least twenty-five centuries, and its
> categories appear (so far) to be incommensurable with or orthogonal to the categories of the
> (neuro-science) whose long-term claim to explain human behavior seems undeniable. Folk
> psychology is nothing more and nothing less than a culturally entrenched theory of how we
> and the higher animals work. It has no [...] special status of any kind whatsoever.[11]

---

[10] Quine paved the way for analytical metaphysics but he himself seldom engaged in *philosophical*
(as opposed to logical) theory construction. His view of mental concepts was that while they were
practically indispensable, there could be no theory of them, scientific or otherwise, for they were
a kind of "dramatic idiom" and "if we are limning the true and ultimate structure of reality, the
canonical scheme for us is the austere scheme that knows [...] only the physical constitution
and behavior of organisms" (Willard van Orman Quine, *Word and Object* [Cambridge: The MIT
Press, 1960]), 221. The latter makes Quine a kind of eliminativist: for him eliminativism is not a
philosophical theory about the mental but the rejection of any such philosophical theory. The aim
of *Word and Object* was not to construct an explanatory philosophical theory of language but to
show the explanatory emptiness of philosophical notions such as meaning, proposition, property,
and the like. Analogous things can be said about Quine's other notable philosophical contributions.
[11] Paul Churchland, "Eliminative Materialism and the Propositional Attitudes," *Journal of Philos-
ophy* 78 (1981), 73.

Fodor, on the other hand, thinks that commonsense psychology is indispensable and can be given a legitimate scientific vindication.

> Vindicating commonsense psychology means showing how you could have [...] a respectable science whose ontology explicitly acknowledges states that exhibit the sorts of properties that common sense attributes to the attitudes. [...] The main thesis of this book is [...] that it is possible to have a scientific psychology that vindicates commonsense belief-desire explanation, (namely) the Representational Theory of Mind.[12]

This disagreement presupposes agreement on the part of both Churchland and Fodor that commonsense psychology stands or falls with whether it can be scientifically vindicated: that is what makes them robust physicalists in the philosophy of mind. They also agree on another point, which is more fundamental though often neglected, namely on what commonsense (or folk) psychology is. In the first instance, the term denotes the conceptual features of our everyday practice of understanding, interpreting, justifying, and explaining ourselves and our actions to each other. We ascribe thoughts, desires, hopes, feelings, and sensations to ourselves and others, it is because we have knowledge of such mental states that we can coordinate our activities with each other, these states play an essential role in our responding to situations in the world as reasons for us to act, and so on. These are things no one would dispute; the question is what to make of them. How should we understand them?

Both Fodor and Churchland think this is not a difficult question – not one that calls for serious philosophical reflection. On their view, commonsense psychology is a theory about the internal causes of our behavior: our beliefs and desires (states, "entities, events, whatever", as Fodor puts it) that have content and causal power and whose interactions are instances of general laws connecting actions with states of belief and desire. Here is how Churchland characterizes commonsense psychology:

> Our commonsense conceptual framework for mental phenomena [is] a theory [...] Each of us understands others, as well as we do, because we share a tacit command of an integrated body of lore concerning the lawlike relations holding among external circumstances, internal states, and overt behavior. [...] This body of lore may quite aptly be called "folk psychology". [...] The recognition that [it] is a theory provides a simple and decisive solution to [...] the problem of other minds. [Belief in other minds] is an *explanatory hypothesis* [...] that provides explanations/predictions/understanding of the individual's continuing behavior. [...] The structural features of folk psychology parallel perfectly those of mathematical physics; the only difference lies in the respective domain of abstract entities they exploit – numbers in the case of physics, and propositions in the case of psychology.[13]

Fodor characterizes it in this way:

> The theory from which we get this extraordinary predictive power is just good old commonsense belief/desire psychology. That's what tells us, for example, how to infer people's intentions from the sounds they make [...] and how to infer people's behavior from their intentions. [...] It takes for granted that overt behavior comes at the end of a causal chain whose links are mental events – hence unobservable – and which may be arbitrarily long. [...] A psychology is commonsensical about the attitudes [...] just in case it postulates

---

[12] Jerry Fodor, *Psychosemantics* (Cambridge: MIT Press, 1987), 3.

[13] Churchland, "Eliminative Materialism and the Propositional Attitudes," 69.

states (entities, events, whatever) satisfying the following conditions: (1) They are semantically evaluable. (2) They have causal powers. (3) The implicit generalizations of commonsense belief/desire psychology are true of them.[14]

Two things about this shared view of commonsense psychology should be noted. The first is that its construal of our everyday practice is badly underargued – indeed, is hasty in the extreme. There are arguments as to why we should call it a theory but they are pursued at a very abstract level, and they simply take it for granted that among the things we have always known are that beliefs and desires are internal states or entities that cause the behavior that is our action. On this view, commonsense psychology is a collection of intuitive judgments, of the things ordinary folk say or would say, of platitudes Granny would deliver – just the sorts of things ordinary language philosophers were criticized for appealing to. No serious attempt is made to comprehend the complexities of our everyday practice and the self-understanding implicit in it, to get some perspicuous overview of it. Real philosophical work, on this view, is not done here but later on.

The second thing to note is that these hasty accounts of our practice are essentially the Cartesian construal of the mental-physical distinction. The physical is regarded as the physicalistic, the mental as internal states and processes that cause behavior physicalistically understood. What is presented as a straight description of phenomena that need scientific vindication and explanation is in fact a theoretically-loaded redescription in Cartesian terms of both the physical and the mental. The physical does not comprise the everyday objects and events we live with, respond to, change and preserve – beautiful and ugly, colored and full of sound – but objects whose properties are definable in physicalistic terms, these other terms applying rather to the internal effects of the physicalistic (which we project back onto the physicalistic). The behavior we observe in each other is physical motion – behavior of organisms at best – but we do not observe people (or animals of any kind) suffering, bored, nervous, angry, satisfied, and so on, for those terms apply not to behavior but to its internal causes.

Commonsense psychology as Fodor and Churchland understand it is, then, a Cartesian theory of the distinction between the physical and the mental, a theory neither challenges but which both take for granted. Where they differ is over what to do at the next stage, where real philosophy, in their view, begins and which involves constructing a theory that will find a place for the mental, understood in the Cartesian way, in a physicalistic view of the world, which essentially means to show how we can be Cartesians without being dualists.

This is relatively easy for Churchland because in describing commonsense psychology, he often expresses the Cartesian distinction itself in terms of his own theory. For example, he accepts the Cartesian view that the asymmetry between the first and third person ascription of mental states should be explained by reference to the distinction between the physical as outer and observable, and the mental as inner and introspectable. But he articulates it in terms of a distinction between an

---

[14] Fodor, *Psychosemantics*, 10, 16.

"explanatory hypothesis" in the case of the mental states of others, and "an acquired habit of conceptual response to one's internal states" in the case of one's own mental states. Fitting *that* distinction into a physicalistic view of the world is not difficult since what it amounts to is accepting the Cartesian distinction but denying that there are any instances on the mental side. The problem, of course, is that there is no longer any reason to accept Churchland's description of commonsense psychology, but from his point of view that does not matter since it is supposed to be eliminated anyway.

Fodor's task is a little harder, for he wants to give a physicalist explanation of the Cartesian distinction that doesn't simply eliminate the mental half. The idea is to be Cartesian without being dualist by taking the mental to consist of mental states that can be incorporated into a scientific theory in the fullest sense of the term. Fodor uses the computational-functionalist theory, which identifies mental states not with physical states but with computational states, which are internal and whose essential (computational) properties are in cyber space and hence unobservable even in principle. The aim is to make the Cartesian distinction scientifically respectable by showing it is isomorphic to the scientific distinction between physical states (hardware) and computational states (software).

Fodor's view cannot fail to impress with its virtuosity and its use of the computer model, which makes it a distinctive twentieth-century contribution to philosophy of mind. It remains wedded to the Cartesian distinction, however, and hence is, as Churchland would put it, "a stagnant or degenerating research program." I will not criticize it here, however, except to offer an account of why it even appears to have some measure of success as an account of the mind.

This has to do with the three conditions Fodor contends must be met by any adequate physicalistic account of the mind – that it postulate states (entities, events, whatever) which are semantically evaluable, have causal powers, and satisfy the generalizations of commonsense belief-desire psychology. What I want to call attention to is what these conditions do *not* include, namely three features that are implicit in Wittgenstein's critique of the Cartesian distinction and which I regard as essential to any adequate account of commonsense psychology.

The first is the normative dimension of commonsense psychology – the fact that our thought and action are *normatively constituted*.[15] To act for a reason is to take some consideration to be, normatively speaking, a reason for one to act, to believe a claim is to take it to be worthy of belief, and to desire something is to take it to be, in some respect, desirable. The second is the centrality of *persons* (or agents). In commonsense psychology, it is persons who believe, desire, hope, feel, suffer, and so on; these states are not internal states, events, or entities, but states of whole persons. Moreover, to understand why people act as they do, we have to grasp what,

---

[15] Fodor's requirement that mental states be "semantically evaluable" has nothing to do with normativity. What he means is that they have truth or reference conditions, neither of which he thinks of as normative even in the broadest sense of that term.

from their point of view, they are doing and why – what they take to be reasons for them to act, not what internal states or events bring about their acting. The third is the irreducibly *interpretive* character of the descriptions and explanations given and received in commonsense psychology. To grasp what people are doing, we have to grasp the reasons for their action, which requires knowing what those actions are, which requires understanding what they are believing or intending, which presumes we understand their behavior, and so on in the "hermeneutic circle". There is no settled starting point for commonsense descriptions and explanations of the actions and mental states of persons: the question of whether someone is in a certain mental state is inseparable from the question of how to describe the state in the first place. There are no standards beyond or beneath these interpretive considerations by which we can judge the interpretive conclusions we reach.

Fodor's construal of commonsense psychology simply omits these essential dimensions, and hence he is not troubled by the fact that they have no place in his scientific explanation of it. The fundamental reason for this omission is that it is intrinsic to the Cartesian distinction itself, even in its dualist version. Normative notions play no constitutive role in the Cartesian account of the mental. Nor do persons as they figure in commonsense psychology play a role, for it is an account of what the mental would be like in isolation from bodily behavior. Finally, neither observation of the physical nor introspection of the mental is interpretive in any significant sense.

The problem with Cartesianism, therefore, is not dualism but the fact that the way it distinguishes the mental from the physical omits these essential features of the mental. Descartes himself did not leave them out in his overall philosophy of mind for they showed up whenever he dealt with the mind–body union. Cartesianism is not Descartes; it is what results when Descartes's dualism is abstracted from his account of the human being as a genuine unity of mind and body. Dropping the dualism from Cartesianism does not change the fundamental point, which is that we have an entirely abstract distinction between the physical and the mental that rules out features that are essential to commonsense psychology.

15.3 Davidson's Philosophy of Mind

I come now to Davidson, my aim being to show why his philosophy of mind should be seen as with and not against Wittgenstein. This may seem surprising, for Davidson's philosophy of mind is often read as making a significant contribution to robust physicalism. He claims that explanations of intentional actions in terms of reasons must be causal and that causation, whether physical or mental, entails that events have physicalistic descriptions, and he accepts the monistic view that every event (and entity) is physical. On this reading, Davidson is a sophisticated contributor to the physicalistic tradition of Carnap and Quine, who shares the standard analytical attitude toward theories of meaning, the mind–body problem, the explanation of action, and so on.

That is the way Davidson is read not only by physicalists who want to enlist him in their cause but also by many Wittgensteinians, who put him near the top of their enemies list. There is no question that Davidson is a monist but the significance of his view is that it rejects both dualism and robust physicalism. I once said to him that I had wondered if he might be a closet Wittgensteinian; his immediate reply was, "Well, I don't know about the closet." I think he spoke the truth if he meant that his work was significantly influenced by Wittgenstein, and I want to show why he should be read that way. I will focus on his philosophy of action because it is basic to all his work.

(1) The center piece of Davidson's philosophy of action is his account of what it is to act for a reason. As is well known, he criticized the Wittgensteinian accounts dominant in the 1950s for having no account of the difference between having a reason to act and acting because of a reason. The difference, he claimed, must be causal: reasons which explain our actions cause them. But his account of what this involves has been read in two ways, one of which yields the main-line causalism robust physicalists hold and the other of which yields a much different view that reflects the fact that the most important influence on his philosophy of action was Elizabeth Anscombe, whose *Intention* he regarded as "the most important treatment of action since Aristotle."[16]

The main-line reading of Davidson is as follows. Reasons explanations are causal, and reasons are beliefs and desires that must, therefore, cause the actions they explain. The traditional view of this was the covering law model, which held that beliefs and desires cause actions in virtue of a causal law that connects descriptions of beliefs and desires *as* reasons with descriptions of actions *as* actions – causal laws, that is to say, at the intentional level. On this reading, Davidson showed two things. First, that there are no such causal laws at the intentional level[17] and, second, that there do not *need* to be any because there are causal laws at the physical level that account for the causal power of reasons. Any belief-desire pair that is the reason for which an agent acts is identical with some physical token, and that physical token has a physical description that is related by causal law to a physical description of the action the reason explains. It is the existence of that causal law at the physical level that distinguishes a reason an agent acts because of from one she merely has. Here is a diagram of this reading.

Belief-desire pair (token *identical* with)	→	(causally *explains*)	→	action (token *identical* with)
physical event	→	(causally *related to*)	→	physical event

---

[16] From the cover of the Harvard edition published in 2000.

[17] "We don't know precise laws for explaining and predicting [psychological events]; but unlike the situation in the natural sciences, this isn't because we haven't discovered them yet; it's because there are no such laws," see Ralf Stoecker (ed.), *Reflecting Davidson* (Berlin: de Gruyter, 1993), 312.

There are very serious problems with such a view[18] that its defenders have tried to deal with by constructing various kinds of theories, but my concern now is whether this is a correct way to read Davidson, which I am convinced that it is not.

Davidson's real view is in his "Actions, Reasons, and Causes" but it is difficult to see there without help from commentators and his later writings. It is this. There is a clear distinction between causal *relations* and causal *explanations*. Causal relations hold only between events, which are entities that can be truly described in any number of ways. They hold no matter how the events are described (sentences ascribing them are *extensional*), but, nevertheless, one event has a causal relation to another only if there is a strict law instantiated by descriptions of the events. This law must be a law of physics (since all strict laws are), and hence any events in causal relations have physical descriptions. It follows that mental events that are causes and effects also have physical descriptions (and are also physical).[19]

Causal *explanations* must describe causes and effects in explanatorily relevant ways, hence their validity depends on how phenomena are described (so that sentences giving a causal *explanation* are intensional). To be causal, an explanation must meet three conditions.

1. It must explain the occurrence of an *event* (or what entails that occurrence). The cause need not be an event; typically it is a state or condition of some kind. Hence, if "A causes B" means that A *causally explains* B, then A need not be an event.
2. It must entail the occurrence of some event *associated with* the cause that is *causally related* to the effect. We need not know which event that is, though it must have a physical description that instantiates a strict law.
3. Its validity must depend on generalizations connecting descriptions of the cause and the effect, which are empirical but are *rough* generalizations and *not* strict laws.

Davidson's claim is that rational explanations meet all three of these conditions. It meets the first because an action is an event. The *cause* that explains an action, however, is not an event, because beliefs and desires are not events but *states* ascribed to whole persons. "'Primary reasons', as I have used the phrase," he writes, "are certainly not events. [...] beliefs and desires are not changes. They are states, and

---

[18] Cf. my "Davidson on Intentional Behavior," in E. Lepore and B. McLaughlin (eds.), *Actions and Events* (Oxford: Basil Blackwell, 1985). I now recognize that I misread Davidson at various points in this paper.

[19] Let me note that does not make Davidson an epiphenomenalist: he does not claim that events are causally related only *in virtue of physical* descriptions or properties. On his view events are not causally related in virtue of anything. Cf. "Thinking Causes," in Heil and Mele (eds.), *Mental Causation* (Oxford: Oxford University Press, 1993), 8, 12: "[...] the efficacy of an event cannot depend on how the event is described, while whether an event can be called mental, or can be said to fall under a law, depends entirely on how the event can be described [...] It is irrelevant to the causal efficacy of physical events that they can be described in the physical vocabulary. It is *events* that have the power to change things, not our various ways of describing them."

since I don't think that states are *entities* of any sort, and so are not events, I do not think beliefs and desires are events."[20]

It meets the second because when an agent acts because of a reason, there is an event *associated with* the reason (a belief-desire pair) that is *causally related* to the action. This event might be almost anything: coming to have a belief, changing one's mind, noticing the corner ahead, or change in the brain, etc. This event does not causally *explain* the action (it is not the reason: that is the belief-desire pair) but is a necessary condition for the explanation to be causal. Nor does it *explain why* the agent acted for one reason rather than another; it is rather part of what it *means* to say he acted for a given reason. As Davidson puts it, "[. . .] The explanation provides no reason for saying that one suitable belief-desire pair rather than another (which may also have been present in the agent) did the causing."[21]

It meets the third condition because its validity depends on empirical generalizations that are not strict laws. Generalizations are necessary because of the dispositional character of beliefs and desires, but these generalizations hold *only* for the particular agent whose action is being explained.[22] If we try to extend them to agents more generally (by adding *ceteris paribus* clauses), they cease to be empirical and become conceptual or normative principles. This means that the real force of a rational explanation turns on the conceptual and normative principles implicit in our *interpretation* of the actions of rational agents in the light of their reasons. Here is a diagram of how I think Davidson should be read.

Belief-desire pair	→	(causally *explains*)	→	action
(*associated* with)				(token *identical* with)
physical event	→	(causally *related to*)	→	physical event[23]

The decisive differences between this reading of Davidson and the previous one are as follows. First, there is no claim that beliefs and desires are physical because Davidson's argument for the identity of the mental and the physical applies only to events, and beliefs and desires are not events (nor entities of any kind). Second, the events that are causally related to the action are not (except occasionally) *reasons* for the action; they may have a connection with the reasons but they need not, and they do not when we do not know what they are.[24] Third, these causes do not explain why

---

[20] *Reflecting Davidson*, 287.

[21] Davidson, "Problems in the Explanation of Action," in *Problems of Rationality* (Oxford: Oxford University Press, 2004), 109.

[22] "The laws implicit in reason explanations are simply the generalizations implied by attributions of dispositions. But then the "laws" are peculiar to individuals at particular moments." (Davidson, "Hempel on Explaining Action," in *Essays on Actions and Events* [Oxford: Oxford University Press, 1980], 265).

[23] Calling the events "physical" is redundant since Davidson holds that to be physical is to have a physical description, which every event does; I call them that because of his claim that if they are causally related, they must have physical descriptions that are an instances of a strict law.

[24] "Sometimes the answer (to "Why did you do it?") will mention a mental event that does not give a reason: "Finally I made up my mind." However, there also seem to be cases of intentional action

an agent acts for one reason and not another: Davidson does not think there is any such explanation. Finally, what in the end give force to a rational explanation are the normative and conceptual principles which are used in an overall interpretation of the agent's action in the light of his attitudes, his situation, and his overall behavior.[25]

(2) Given this reading of Davidson's account of rational explanation, we can see the ways in which his point of view in the philosophy of mind is in all major respects *with* Wittgenstein and *against* robust physicalism. The most significant of these is Davidson's clear-headed rejection of the Cartesian distinction between the mental and the physical.

"There are no such things as minds, but people have mental properties, which is to say that certain psychological predicates are true of them."[26] That is Davidson's account of mental states: they are not internal entities – indeed, they are not entities at all – but states of whole persons. To ascribe a belief to a person is to describe *her* in a certain way – to apply a mental predicate to *her* – which is to say that *she* is in a certain mental state. Mental states differ from physical states in that we ascribe them to ourselves without evidence or inference, but this is not because mental states are internal while physical states are external, or because mental states must be introspected by their bearers. It is because to determine the mental states of others, I must interpret their actions and speech whereas it makes no sense to speak of interpreting myself.[27]

Interpretation has no significant role to play in physicalist philosophy of mind because it has no significant role to play in the Cartesian distinction between the physical and the mental. It is, however, at the center of Davidson's philosophy of mind. He writes that "What a fully informed interpreter could learn about what a speaker means is all there is to learn; the same goes for what the speaker believes."[28] Intentional states simply are what must be ascribed to an agent to render her behavior (including her speech) intelligible to an interpreter. They cannot, therefore, be *entities* of any kind because, while there are endless truths about entities, intentional

---

where we cannot explain at all why we acted when we did. In such cases, explanation in terms of primary reasons parallels the explanation of the collapse of the bridge from a structural defect: we are ignorant of the event or sequence of events that led up to (caused) the collapse, but we are sure that there was such an event or sequence of events." Davidson, "Actions, Reasons and Causes," 13.

[25] For further discussion and defense of this way of reading Davidson, see my "Intentionalists and Davidson on Rational Explanation," in G. Meggle (ed.), *Actions, Norms and Values* (Walter de Gruyter, 1991), 191–208.

[26] Davidson, "Donald Davidson," in Samuel Guttenplan (ed.), *A Companion to the Philosophy of Mind* (Oxford: Blackwell, 1994).

[27] "The existence of first person authority is not an empirical discovery, but rather a criterion, among others, of what a mental state is. [. . .] Exceptions do not throw in doubt the presumption that we know our own minds. What accounts for this presumption? [. . .] [The answer is that] we must interpret the thoughts of others on the basis of evidence; interpreting ourselves does not [. . .] make sense [. . .]. The difference [. . .] is that when I interpret you, two languages are involved, yours and mine (the same words may mean different things in your language and mine). In the second case, only one language is involved, my own; interpretation is therefore not (exceptional cases aside) in the picture." "Donald Davidson," in *A Companion to the Philosophy of Mind*.

[28] Davidson, "A Coherence Theory of Truth and Knowledge," in *Subjective, Intersubjective, Objective* (Oxford: Oxford University Press, 2001), 148.

states are just what an adequate interpretation takes them to be, and hence there cannot be truths about them only scientific experts (like neuro- or information-scientists) can establish.[29]

As we have seen, Davidson elucidates the asymmetry between ascribing mental states to others and to oneself by saying we interpret others but it makes no sense to interpret ourselves. He develops this point further in contending that interpretation is the "source of the ultimate difference" between the mental and the physical, a point he develops in the following striking account of interpretation.

> Success in interpretation is always a matter of degree: the resources of thought or expression available to an interpreter can never perfectly match the resources of the interpreted. We do the best we can. [. . .] This is the process of radical interpretation. There is no further court of appeal, no impersonal objective standard against which to measure our own best judgments of the rational and the true.
>
>   Here lies the source of the ultimate difference between the concepts we use to describe mental events and the concepts we use to describe physical events, the difference that rules out the existence of strict psychophysical laws. The physical world and the numbers we use to calibrate it are common property, the material and abstract objects and events that we agree on and share. But it makes no sense to speak of comparing, or coming to agree on, ultimate common standards of rationality, since it is our own standards in each case to which we must turn in interpreting others. This should not be thought of as a failure of objectivity, but rather as the point at which 'questions come to an end.' Understanding the mental states of others and understanding nature are case where questions come to an end at different stages. How we measure physical quantities is decided intersubjectively. We cannot in the same way go behind our own ultimate norms of rationality in interpreting others. Priority is not an issue. We would have no full-fledged thoughts if we were not in communication with others, and therefore no thoughts about nature: communication requires that we succeed in finding something like our own patterns of thought in others.[30]

The two other essential features of commonsense psychology that I noted as absent from the Cartesian distinction are implicit in this notion of interpretation. That the mental is normatively constituted is a prominent theme in Davidson's work, in particular in his arguments that psychology is autonomous and not dependent on the natural sciences. "The study of human action, motives, desires, beliefs, memory, and learning," he writes, "cannot employ the same methods as, or be reduced to, the more precise physical sciences."[31]

> The reason mental concepts cannot be reduced to physical concepts is the *normative* character of mental concepts. [. . .] The semantic contents of attitudes and beliefs determine their relations to one another and to the world in ways that meet at least rough standards of consistency and correctness. Unless such standards are met to an adequate degree, nothing can count as being a belief, a pro-attitude, or an intention. But these standards are norms – *our* norms – there being no others.[32]

---

[29] "If you ask what kind of properties we're attributing when we attribute beliefs, I think a theory about how we tell that belief-attributing sentences are true provides the best answer. This shows what kind of property it is: it's a property which you determine to apply to an individual in the following way. [. . .] (and here you describe the method). Is there something more to say about it? I don't see why there has to be." *Reflecting Davidson*, 288.

[30] Davidson, "Donald Davidson," in *A Companion to the Philosophy of Mind*, 232

[31] Davidson, "Comments and Replies," in *Essays on Actions and Events*, 240.

[32] Davidson, "Problems in the Explanation of Action," 114.

Davidson also presumes the centrality of persons in commonsense psychology. He regards mental states as states of whole persons, and his way of understanding thought, language and action never loses sight of thinkers, speakers, and agents. He further maintains that to understand mental states, we have to be able to grasp the point of view of their possessors. This is what the principle of charity is all about: understanding another person requires a large enough measure of agreement to ensure that we share a world and a point of view on that world. That does not rule out disagreement; the point is that disagreement presumes a large measure of agreement. "Some kind of basic agreement, not just in beliefs, but also in values, is essential to understanding. [. . .] Understanding another person depends upon finding common grounds not only with respect to beliefs but also with respect to values, right from the start."[33] Reason-explanations, he claims in another place, "make others intelligible to us only to the extent that we can recognize something like our own reasoning powers at work, (powers which cannot) be reduced to non-normative, perhaps formal, characteristics."[34]

Davidson's conception of the mental–physical distinction also rejects the Cartesian conception of the physical as merely physicalistic. What the physical is, is what it presents itself as to ordinary observation, and we can grasp the mental only as we grasp the physical. "Only those who share a common world can communicate; only those who communicate can have the concept of an intersubjective, objective world."[35] There are non-interpretive truths about the physical world – that is a distinguishing mark of the physical – but only those capable of interpretation can grasp or establish them.

In all these ways, then, Davidson's philosophy of mind is anti-Cartesian, and hence anti-physicalist, its main themes being very much in the spirit of Wittgenstein. The same is true, I would argue, of his conception of philosophy. To many, Davidson's way of working seems opposed to Wittgenstein's because of the extensive use he makes of the notion of theory – a theory of meaning, a theory of interpretation, a theory of the attitudes, and so on. But the fact that a philosopher characterizes his work in terms of "theory" does not mean he is constructing the kind of theories Wittgenstein urged philosophers to avoid. What Wittgenstein rejected were theories that, on the one hand, obscured or distorted our everyday practice of describing and explaining ourselves and our world and, on the other hand, tried to vindicate and explain that practice by showing why it *must* be so and so or how it can be altered to meet external standards, scientific or otherwise.

Davidson's theories do not do that. His theory of meaning, for example, is a real theory in the technical sense of having axioms and theorems, but it is not *explanatory*. It does not undertake to explain why we speak as we do, why our language must have certain forms, what understanding a language consists in, etc. It

---

[33] *Ibid.*, 115.

[34] Lars Bergström and Dagfinn Føllesdal, "Interview with Donald Davidson in November 1993," *Theoria* 60 (1994), 210.

[35] Davidson, "Donald Davidson," in *A Companion to the Philosophy of Mind*, 234

describes a certain dimension of our linguistic practice in a systematic way that aims at being sensitive to that practice, something that led Davidson to speak in terms of a "passing theory" that applies only to a particular communicative situation. Such a systematic description of our linguistic practice may be objectionable in a number of ways, but it does not violate the spirit of Wittgenstein's conception of philosophy, or of language as "woven" into actions. "The concept of meaning", Davidson wrote, "derives all of its content from the case of successful interpretation. That is, cases where a person intends to be interpreted in a certain way and is."[36]

Davidson's use of "theory" in such contexts as "theory of interpretation" or "theory about how we tell that belief-attributing sentences are true" is not a technical use. It simply means an account that is discursive, articulate, systematic, rigorous, and the like, and has nothing to do with scientific theories. Davidson did think it was possible to do philosophy with the same measure of care, explicitness, and argument that characterizes the best kind of scientific work, but that is not to confuse philosophy with scientific theorizing. He showed only a passing interest in the metaphysical theorizing that has taken over analytical philosophy, and even his work on events, identity, and causation was pursued in the context of concrete examples and in relation to problems that arise from the subject matter and not from theories about it. He pursued supervenience only in an informal way and was never tempted by the view that it *explains* anything. His conception of truth avoided the notion that sentences are true *in virtue of* anything, and he offered no theory of truth. "Fact," he wrote, "is for me [. . .] just a general word and you can say there are moral facts as well as others,"[37] and as for properties, they are simply predicates true of something.

Davidson's aim can be characterized as giving a perspicuous representation of what we already know. He often did it in ways Wittgenstein did not, but to learn from someone of Wittgenstein's stature requires not following him slavishly or in a spirit of partisanship. It has been a long struggle to save St. Thomas from the Thomists, Descartes from the Cartesians, Dewey from the Deweyans. Now the struggle should be to save Wittgenstein from the Wittgensteinians, and there is no better way to engage in that struggle than to learn from Davidson.

---

[36] "Interview with Donald Davidson . . .," *Theoria*, 1994.

[37] *Ibid.*

Bibliography

Achillini, Alessandro, see Alessandro Achillini.

Agostino Nifo. *Expositio subtilissima in tres libros Aristotelis De anima*. Venice, 1559.

Agostino Nifo. *Expositio subtilissima collectanea commentariaque in III libros Aristotelis De anima.* Venice, 1553.

Alanen, Lilli. *Descartes' Concept of Mind*. Cambridge: Harvard University Press, 2003.

Albert the Great (Albertus Magnus). *De anima*. Edited by Clemens Stroick. Aschendorff, 1968.

Alessandro Achillini. *Quaestiones quodlibeta de intelligentiis*. Bologna, 1494.

Alexander Aphrodisias. *The* De anima of *Alexander of Aphrodisias*. Translated and commented by Athanasios P. Fotinis. Washington, D.C.: University Press of America, 1979.

Allison, Henry E. *Kant's Theory of Freedom*. Cambridge: Cambridge University Press, 1990.

Allison, Henry E. *Kant's Transcendental Idealism*: A*n Interpretation and Defence*. Revised & Enlarged Edition. New Haven: Yale University Press, 2004.

Alvarez, Balthasar. *Tractatus de anima separata*. Appended in *Commentarii Collegii Conimbricensis Societatis Jesu in tres libros De anima Aristotelis Stagiritae.* Lyon 1600.

Ameriks, Karl. *Kant's Theory of Mind*. Oxford: Clarendon Press, 1982.

Antonio Trombetta. *Tractatus singularis contra Averroistas de humanarum animarum plurificatione ad catholicae fidei obsequium Patavii editus*. Venice, 1498.

Aphrodisias, Alexander, see Alexander Aphrodisias.

Aquinas, Thomas, see Thomas Aquinas.

Aristotle. *The Complete Works of Aristotle*, Vol. 1 and 2. Edited by Jonathan Barnes. Princeton: Princeton University Press, 1984.

Aristotle. *On the Soul*. Translated by W.S. Hett. Loeb Classical Library. Cambridge, Mass.: Harvard University Press, 1957.

Azouvi, François. *Maine de Biran: la science de l'homme*. Paris: Vrin, 1995.

Baertschi, Bernard. *Les rapports de l'âme et du corps: Descartes, Diderot et Maine de Biran*. Paris: J. Vrin, 1992.

Bartholomaeus Arnoldi de Usingen. *Exercitium de anima*. Erfurt: Wolffgang Schenck, 1507.

Bartholomaeus Arnoldi de Usingen. *Compendium naturalis philosophie*. Erfurt: Wolffgang Schenk, 1505–1507.

Bartholomaeus Arnoldi de Usingen. *Parvulus philosophie naturalis*. Leipzig: Wolfgang Stöckel, 1499.

Beiser, F. "Kant's Intellectual Development: 1746–1781." In *The Cambridge Companion to Kant*. Edited by Paul Guyer. Cambridge: Cambridge University Press, 26–61, 1992.

Bergson, Henri. *Essai sur les données immédiates de la conscience*. Paris: Alcan, 1889.

Bergson, Henri. *Introduction à la métaphysique* [1903]. In *Œuvres. Edition du Centenaire*. Paris: PUF, 1991.

Bergson, Henri. *Matter and Memory*. Translated by N.M. Paul and W.S. Palmer. New York: Zone Books, 1990.

Bergson, Henri. *Time and Free Will: An Essay on the Immediate Data of Consciousness*. Translated by F.L. Pogson. Montana: Kessinger Publishing Company, 1910.

Bergström, Lars and Føllesdal, Dagfinn. "Interview with Donald Davidson in November 1993." *Theoria* 60: 207–25, 1994.

Bermudez, Jose Louis, Marcel, Anthony and Eilan, Naomi. *The Body and the Self*. Cambridge: MIT Press, 1995.

Bode, Boyd H. " 'Pure Experience' and the External World." *The Journal of Philosophy* 2: 128–33, 1905.

Bourne, R.E. *Thomistic Psychology*. New York: Macmillan, 1957.

Brentano, Franz. *Alte und neue Logik* (manuscript EL 108/2, 1877). Graz: Forschungsstelle und Dokumentationszentrum für österreichische Philosophie.

Brentano, Franz. *Descriptive Psychology*. Edited and translated by Benito Müller. London and New York: Routledge, 1995.

Brentano, Franz. *Kategorienlehre*. Edited by A. Kastil. Leibzig: Meiner, 1933.

Brentano, Franz. *Die Psychologie des Aristoteles, insbesondere seine Lehre vom* nous poiētikos. Mainz: Kirchheim, 1867; Repr. Darmstadt: Wissenschaftliche Buchgesellschaft, 1960.

Brentano, Franz. *Psychology from an Empirical Standpoint*. Translated by Antos C. Rancurello, P.B. Terrell, and Linda L. McAlister. London: Routledge, 1995 [1874].

Brentano, Franz. *Unsterblichkeit* (manuscript LS 22/1, 1875/76). Graz: Forschungsstelle und Dokumentationszentrum für österreichische Philosophie.

Brentano, Franz. *Über Aristoteles: Nachgelassene Aufsätze*. Edited by Rolf George. Hamburg: Meiner, 1986.

Brook, Andrew. *Kant and the Mind*. Cambridge: Cambridge University Press, 1994.

Brooks, G.P. and Johnson, R.W. "Johann Caspar Lavater's *Essays on physiognomy*." *Psychological Reports* 46: 16–9, 1980.

Buridan, John, see John Buridan.

Cajetanus, Thomas, see Thomas Cajetanus.

Callergård, Robert. *An Essay on Thomas Reid's Philosophy of Science*. Acta Universitatis Stockholmiensis, Stockholm Studies in Philosophy 28, 2006.

Campbell, John. *Reference and Consciousness*. Oxford: Clarendon Press, 2002.

Carl, Wolfgang. *Der Schweigende Kant*: *Die Entwürfe zu einer Deduktion der Kategorien*. Göttingen: Vandenhoeck & Ruprecht, 1989.

Carnap, Rudolf. "Psychologie in physikalischer Sprache." *Erkenntnis* 3 (1932): 107–42. Translated by G. Schick: "Psychology in Physical Language." In *Logical Positivism*. Edited by A. J. Ayer. New York: Free Press, 165–97, 1959.

Carr, David. *The Paradox of Subjectivity: The Self in the Transcendental Tradition*. Oxford: Oxford University Press, 1999.

Cassirer, Ernst. *Kant's Life and Thought*. New Haven and London: Yale University Press, 1981.

Castellote Cubells, Salvador. *Die Anthropologie des Suarez: Beiträge zur spanischen Anthropologie des XVI. und XVII. Jahrhunderts*. Freiburg: Karl Alber, 1962.

Chisholm, Roderick. *The Foundations of Knowing*. Minneapolis: University of Minnesota Press, 1981.

Chrysostomus, Javelli. *Quaestiones naturales super tres libros Aristotelis De anima. Opera* I. Lyon, 1580.

Churchland, Patricia S. *Neurophilosophy: Toward a Unified Science of the Mind-Brain*. Cambridge: The MIT Press, 1986.

Churchland, Paul. "Eliminative Materialism and the Propositional Attitudes." *Journal of Philosophy* 78: 67–90, 1981.

Commentarii Collegii Conimbricensis Societatis Jesu in tres libros De anima Aristotelis Stagiritae. Lyon, 1600.

Comte, Auguste. *Cours de philosophie positive*. Paris: Rouen frères, 1830; Repr., Paris: Hermann, 1975.

Cooper, David E. *The Measure of Things: Humanism, Humility, and Mystery*. Oxford: Clarendon Press, 2002.

Cormier, Harvey. *The Truth Is What Works: James, Peirce, and the Seeds of Death*. Lanham, MD: Rowman & Littlefield, 2001.

Cottingham, John. "Descartes' Treatment of Animals." In *Descartes*. Oxford Readings in Philosophy. Edited by J. Cottingham. Oxford: Oxford University Press, 225–33, 1998.

Cournot, Antoine August. *Essai sur les fondements de nos connaissances et sur les caractères de la critique philosophique*. Paris: Vrin-CNRS, 1975 [1851].

Crane, Tim. "Introspection, Intentionality, and the Transparency of Experience." *Philosophical Topics* 28: 49–67, 2000.

Crane, Tim and Patterson, Sarah. *History of the Mind–Body Problem*. London: Routledge, 2000.

Cranz, Edward F. "The Publishing History of the Aristotle Commentaries of Thomas Aquinas." *Traditio* 34: 157–92, 1978.

Crowell, Steven Galt. "Facticity and Transcendental Philosophy." In *From Kant to Davidson: Philosophy and the Idea of the Transcendental*. Edited by Jeff Malpas. London: Routledge, 13–29, 2002.

Crowell, Steven Galt. *Husserl, Heidegger, and the Space of Meaning*. Evanston: Northwestern University Press, 2001.

Davidson, Donald. "Donald Davidson." In *A Companion to the Philosophy of Mind*. Edited by Samuel Guttenplan. Oxford: Blackwell, 231–36, 1994.

Davidson, Donald. *Essays on Actions and Events*. Oxford: Oxford University Press, 1980.

Davidson, Donald. *Problems of Rationality*. Oxford: Oxford University Press, 2004.

Davidson, Donald. *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press, 2001.

Davidson, Donald. "Thinking Causes." In *Mental Causation*. Edited by Heil and Mele. Oxford: Oxford University Press, 1993.

de Biran, François Marie Maine. *Essai sur les fondements de la psychologie* (1812). In *Œuvres*. Edited by F.C.T. Moore. Paris: Vrin, 2001.

de Biran, Francois Marie Maine. *Nouveaux essais d'anthropologie. Note sur l'idée d'existence. Derniers fragments*. In *Œuvres*. Edited by Bernard Baertschi. Paris: J. Vrin 1989 [1824].

de Biran, Francois Marie Maine. *Rapports du physique et du moral chez l'homme*. In *Œuvres*. Edited by F.C.T. Moore. Paris: J. Vrin, 1984.

de Lutrea, Johannes Carnificis, see Johannes Carnificis de Lutrea.

de Usingen, Bartholomaeus, see Bartholomaeus Arnoldi de Usingen.

De Paul, Michael R. and Ramsey, William. *Rethinking Intuition*. Lanahm: Rowman and Littlefield, 1998.

DeArmey, Michael H. "William James and the Problem of Other Minds." *The Southern Journal of Philosophy* 20: 325–36, 1982.

Dennett, Daniel. *The Intentional Stance*. Cambridge: The MIT Press, 1987.

Des Chenes, Dennis. *Life's Form: Late Aristotelian Conceptions of the Soul*. Ithaca: Cornell University Press, 2000.

Descartes, René. *The Philosophical Writings of Descartes I–III*. Translated by J. Cottingham, R. Stoothof & D. Murdoch. Cambridge: Cambridge University Press, 1984–1994.

Descartes, René. *Principles of Philosophy*. Translated by V.R. Miller and R.P. Miller. Dordrecht: Reidel, 1983.

Descartes, René. *Œuvres de Descartes I–XI*. Edited by Charles Adam and Paul Tannery. Paris: J. Vrin, 1964–1976.

Dewey, John. *The Quest for Certainty: A Study of the Relation between Knowledge and Action*. New York: G.P. Putnam's Sons, 1960 [1929].

Dillon, Martin C. "Gestalt Theory and Merleau-Ponty's Concept of Intentionality." *Man and World* 4: 436–59, 1971.

Dillon, Martin C. *Merleau-Ponty's Ontology*. Bloomington and Indianapolis: Indiana University Press, 1988.

Dretske, Fred. *Explaining Behavior: Reasons in a World of Causes*. Cambridge: The MIT Press, 1988.

Dretske, Fred. *Knowledge and the Flow of Information*. Cambridge: The MIT Press, 1981.

Dretske, Fred. "Misrepresentation." In *Belief*. Edited by R. Bogdan. Oxford: Oxford University Press, 17–36, 1986.

Dretske, Fred. *Naturalizing the Mind*. Cambridge: The MIT Press, 1995.

Dummett, Michael. *Origins of Analytic Philosophy*. London: Duckworth, 1993.

Duns Scotus, John, see John Duns Scotus.

Eck, John, see John Eck (Johannes Eckius).

Ehrenfels, Christian von. "On 'Gestalt Qualities'." In *Foundations of Gestalt Theory*. Edited by Barry Smith. München, Wien: Philosophia Verlag, 82–123, 1988.

Embree, Lester. "Biographical Sketch of Aron Gurwitsch." In *Life-World and Consciousness*. Edited by Lester Embree. Evanston, Illinois: Northwestern University Press, 18–30, 1972.

Embree, Lester. "Gestalt Psychology." In *Encyclopedia of Phenomenology*. Edited by Lester Embree, Elisabeth A. Behnke, David Carr et al. Dordrecht: Kluwer, 276–81, 1997.

Embree, Lester. "Gurwitsch's Critique of Merleau-Ponty." *Journal of the British Society for Phenomenology* 12: 151–63, 1981.

Engel, Pascal. "Lagneau entre Ribot et Bergson." In *Jules Lagneau: cours integral 1886–7*. Dijon: CRDP de Bourgogne, 90–100, 1997.

Engel, Pascal. "Plenitude and contingency: modal concepts in 19th century French philosophy." In *Modern Modalities*. Edited by Simo Knuuttila. Dordrech: Kluwer, 179–237, 1988.

Engel-Tiercelin, Claudine. "Reid and Peirce on Belief." In *The Philosophy of Thomas Reid*. Edited by M. Dalgarno and E. Matthews. Dordrecht: Kluwer Academic Publishers, 1989.

Fancher, Raymond. *Pioneers of Psychology*. New York: Norton, 1979.

Fechner, Gustav Theodor. *Elemente der Psychophysik*. Leipzig: Breitkopf & Härtel, 1860.

Ferrariensis, Franciscus Silvester, see Franciscus Silvester Ferrariensis.

Flanagan, Owen. *The Science of the Mind*. Cambridge: The MIT Press, 1991.

Fodor, Jerry. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge: The MIT Press, 1987.

Foucault, Michel. *Les mots et les choses*. Paris: Gallimard, 1966.

Foucault, Michel. "Nietzsche, la généalogie, l'histoire." In *Dits et écrits 1954–1988, I: 1970–1975*. Edited by Daniel Defert and François Ewald with Jacques Lagrange. Paris: Gallimard, 136–56, 1998. An English translation "Nietzsche, Genealogy, History." Translated by J. Harari. In *The Foucault Reader*. Edited by Paul Rabinow. Harmondsworth: Penguin, 76–100, 1984.

Franciscus Silvester Ferrariensis. *Quaestiones eruditissimae in tres libros De anima Aristotelis Stagiritae*. Rome, 1577.

Frederici Vescovini, Graziella. *Le Quaestiones de anima di Biagio Pelacani da Parma*. Florence: Leo S. Olschki Editore, 1974.

Fugali, Edoardo. *Die Zeit des Selbst und die Zeit des Seienden: Bewusstsein und innerer Sinn in Franz Brentano*. Würzburg: Königshausen & Neumann, 2004.

Gaita, Raimond. *Good and Evil: An Absolute Conception*. London and New York: Routledge, 2004.

Gale, Richard M. *The Divided Self of William James*. Cambridge: Cambridge University Press, 1999.

Gale, Richard M. *The Philosophy of William James: An Introduction*. Cambridge: Cambridge University Press, 2005.

Gallagher, Shaun and Cole, Jonathan. "Body Image and Body Schema." In *Body and Flesh: A Philosophical Reader*. Edited by Donn Welton. Oxford: Blackwell, 136–47, 1998.

Gaukroger, Stephen. *Descartes: An Intellectual Biography*. Oxford: Oxford University Press, 1995.

Gendler, Tamar S. and Hawthorne, John. *Conceivability and Possibility*. Oxford: Oxford University Press, 13–26, 2002.

George, Rolf. "Brentano's relations to Aristotle." *Grazer Philosophische Studien* 5: 242–54, 1978.

Giles of Rome (Aegidius Romanus). *Commentationes physicae et metaphysicae in De anima libros tres*. Oberursel (Ursellis), 1604.

Gilson, Étienne. "Cajétan et l'humanisme théologique." *Archives d'histoire doctrinale et littéraire du moyen âge* 30: 113–36, 1955.

Godwin, William. *Memoirs of the Author of "The Rights of Woman"*. Edited by Richard Holmes. London: Penguin Books, 1987.

Goethe, Johann Wolfgang. *Dichtung und Wahrheit*. Goethe Werke, vol. 5. Edited by Klaus-Detlef Müller. Darmstadt: Wissenschaftliche Buchgesellschaft, 1998.

Goldman, Alvin. *Epistemology and Cognition*. Cambridge: Harvard University Press, 1986.

Goldman, Alvin. *Philosophical Applications of Cognitive Science*. Colorado: Westview Press, 1998.

Goodman, Russell B. *Wittgenstein and William James*. Cambridge: Cambridge University Press, 2002.

Gracia, Jorge. "Suárez (and Later Scholasticism)." In *Routledge History of Philosophy III: Medieval Philosophy*. Edited by J. Marenbon. London: Routledge, 452–74, 1998.

Graham, John. "Lavater's *Physiognomy* in England." *Journal of the History of Ideas* 22/4: 561–72, 1961.

Gramsch, Robert. *Erfurter Juristen im Spätmittelalter*. Leiden: Brill, 2003.

Grendler, Paul F. *The Universities of the Italian Renaissance*. Baltimore and London: The Johns Hopkins University Press, 2002.

Griffin, David Ray. *Religion and Scientific Naturalism: Overcoming the Conflicts*. Albany: SUNY Press, 2000.

Gueroult, Martial. *Spinoza*. Hildesheim: Olms, 1974.

Gurwitsch, Aron. "Critical Study of Husserl's *Nachwort*." Translated by Frederick Kersten. In *Studies in Phenomenology and Psychology*. Evanston: Northwestern University Press, 107–15, 1966.

Gurwitsch, Aron. "Edmund Husserl's Conception of Phenomenological Psychology." In *Phenomenology and the Theory of Science*. Edited by Lester Embree. Evanston, Illinois: Northwestern University Press, 77–112, 1974.

Gurwitsch, Aron. *Field of Consciousness*. Pittsburg: Duquesne University Press, 1964.

Gurwitsch, Aron. "Phänomenologie der Thematik und des reinen Ich: Studien über Beziehungen von Gestalttheorie und Phänomenologie." *Psychologische Forschung*, 12: 279–383, 1929. An English translation "Phenomenology of Thematics and of the Pure Ego: Studies of the Relation between Gestalt Theory and Phenomenology." In *Studies in Phenomenology and Psychology*. Evanston, Illinois Northwestern University Press, 175–286, 1966.

Gurwitsch, Aron. "Review of *Phénoménologie de la perception*." *Philosophy and Phenomenological Research* 10: 442–45, 1950.

Gurwitsch, Aron. *Studies in Phenomenology and Psychology*. Evanston: Northwestern University Press, 1966.

Guyer, Paul. *Kant and the Claims of Knowledge*. New York: Cambridge University Press, 1988.

Guyer, Paul. "Kant on Apperception and A Priori Synthesis." *American Philosophical Quarterly* 17: 205–12, 1980.

Hand, Sean (ed.). *The Levinas Reader*. Oxford: Blackwell, 1989.

Harman, Gilbert. "The Intrinsic Quality of Experience." In *Philosophical Perspectives 4: Action Theory and Philosophy of Mind*. Edited by J.E. Tomberlin. Atascadero: Ridgeview, 31–52, 1990.

Harper, William. "Idea." In *A Companion to Metaphysics*. Edited by J. Kim and E. Sosa. Oxford: Oxford University Press, 1995.

Harper, William. "Newton's Argument for Universal Gravitation." In *The Cambridge Companion to Newton*. Edited by I. Bernard Cohen and George E. Smith. Cambridge: Cambridge University Press, 174–201, 2002.

Hatfield, Gary. "Animals." In *Companion to Descartes*. Edited by J. Carriero and J. Broughton. Oxford: Blackwell, 404–25, 2007.

Hatfield, Gary. "Behaviorism and Psychology." In *Cambridge History of Philosophy, 1870–1945*. Edited by Thomas Baldwin. Cambridge: Cambridge University Press, 640–48, 2003.

Hatfield, Gary. *Descartes and the Meditations*. London: Routledge, 2003.

Hatfield, Gary. "Descartes' Naturalism about the Mental." In *Descartes' Natural Philosophy*. Edited by Stephen Gaukroger, John Schuster, and John Sutton. London: Routledge, 630–58, 2000.

Hatfield, Gary. "Descartes' Physiology and Its Relation to His Psychology." In *Cambridge Companion to Descartes*. Edited by John Cottingham. Cambridge: Cambridge University Press, 335–70, 1992.

Hatfield, Gary. "Did Descartes Have a Jamesian Theory of the Emotions?" *Philosophical Psychology* 20: 413–40, 2007.

Hatfield, Gary. "Empirical, Rational, and Transcendental Psychology: Psychology as Science and as Philosophy." In *Cambridge Companion to Kant*. Edited by Paul Guyer. Cambridge: Cambridge University Press, 200–27, 1992.

Hatfield, Gary. "Epistemology and Science in the Image of Modern Philosophy: Rorty on Descartes and Locke." In *Future Pasts: Reflections on the History and Nature of Analytic Philosophy*. Edited by Juliet Floyd and Sanford Shieh. New York: Oxford University Press, 393–413, 2001.

Hatfield, Gary. "Introspective Evidence in Psychology." In *Scientific Evidence: Philosophical Theories and Applications*. Edited by Peter Achinstein. Baltimore: Johns Hopkins University Press, 259–86, 2005.

Hatfield, Gary. *The Natural and the Normative: Theories of Spatial Perception from Kant to Helmholtz*. Cambridge: The MIT Press, 1990.

Hatfield, Gary. "Objectivity and Subjectivity Revisited: Color as a Psychobiological Property." In *Colour Vision: From Light to Object*. Edited by Rainer Mausfeld and Dieter Heyer. Oxford: Oxford University Press, 187–202, 2003.

Hatfield, Gary. "The *Passions of the Soul* and Descartes' Machine Psychology." *Studies in History and Philosophy of Science* 38: 1–35, 2007.

Hatfield, Gary. "The Reality of Qualia." *Erkenntnis* 66: 133–68, 2007.

Hatfield, Gary. "Remaking the Science of Mind: Psychology as a Natural Science." In *Inventing Human Science*. Edited by Christopher Fox, Roy Porter, and Robert Wokler. Berkeley: University of California Press, 184–231, 1995.

Hatfield, Gary. "Sense-Data and the Mind–Body Problem." In *Perception and Reality: From Descartes to the Present*. Edited by Ralph Schumacher. Berlin: Mentis Verlag, 305–31, 2004.

Hatfield, Gary. "The Workings of the Intellect: Mind and Psychology." In *Logic and the Workings of the Mind: The Logic of Ideas and Faculty Psychology in Early Modern Philosophy*. Edited by Patricia Easton. North American Kant Society Publications vol. 5. Atascadero: Ridgeview Publishing Co., 21–45, 1997.

Hatfield, Gary and Epstein, William. "The Sensory Core and the Medieval Foundations of Early Modern Perceptual Theory." *Isis* 70: 363–84, 1979.

Heidegger, Martin. *Grundprobleme der Phänomenologie (1919/1920)*. Gesamtausgabe Band 58. Frankfurt am Main: Vittorio Klostermann, 1993.

Heidegger, Martin. *Sein und Zeit*. Tübingen: Max Niemeyer, 1986.

Heinämaa, Sara. "From Decisions to Passions." In *Merleau-Ponty's Reading of Husserl*. Edited by Ted Toadvine and Lester Embree. Dordrecht, Boston, London: Kluwer, 127–46, 2002.

Heinämaa, Sara. "The Living Body and its Position in Metaphysics: Merleau-Ponty's Dialogue with Descartes." In *Metaphysics, Facticity, Interpretation*. Edited by Dan Zahavi, Sara Heinämaa and Hans Ruin. Dordrecht, Boston, London: Kluwer, 23–48, 2003.

Heinämaa, Sara. "Personality, Anonymity and Sexual Difference." In *Time in Feminist Philosophy*. Edited by Helen Fielding, Dorothea Olkowski and Christina Schües. Indiana University Press, forthcoming in 2009.

Heinämaa, Sara. "Selfhood, Consciousness, and Embodiment: A Husserlian Approach." In *Consciousness: From Perception to Reflection in the History of Philosophy*. Edited by Sara Heinämaa, Pauliina Remes and Vili Lähteenmäki. Dordrecht: Springer, 311–28, 2007.

Heinämaa, Sara. *Toward a Phenomenology of Sexual Difference: Husserl, Merleau-Ponty, Beauvoir*. Lanham, Boulder, New York, Oxford: Rowman & Littlefield, 2003.

Heinämaa, Sara, Lähteenmäki, Vili, and Remes, Pauliina (eds.). *Consciousness: From Perception to Reflection*. Dordrecht: Springer, 2007.

Henry, Michel. *Philosophie et phénoménologie du corps*. Paris: PUF, 1967.

Hesse, Mary B. *Forces and Fields*. London: T. Nelson, 1961.

Hintikka, Jaakko. "Concepts of Scientific Method from Aristotle to Newton." In *Knowledge and the Sciences in Medieval Philosophy*. Acta Philosophica Fennica, Vol. 48. Edited by Monika Asztalos, John E. Murdoch, and Ilkka Niiniluoto. Helsinki, 72–84, 1990.

Hintikka, Jaakko and Remes, Unto. *The Method of Analysis*. Dordrecht: D. Reidel, 1974.

Hobbes, Thomas. *The Questions Concerning Liberty, Necessity and Chance, clearly stated between Dr Bramhall Bishop of Derry, and Thomas Hobbes of Malmesbury*. London, 1656.

Hoenen, Maarten J. F. M. "Die Intellektlehre des Johannes Buridan – ihre Quellen und historisch-doktrinären Zügen." In *John Buridan: A Master of Arts. Some Aspects of his Philosophy*. Aristarium Supplementa VIII. Edited by E. P. Bos and H. A. Krop. Nijmegen: Ingenium, 89–106, 1993.

Hoenen, Maarten J.F.M. "*Via antiqua* and *via moderna* in the Fifteenth Century: Doctrinal, Institutional, and Church Political Factors in the *Wegestreit*." In *The Medieval Heritage in Early Modern Metaphysics and Modal Theory*. Edited by L. Nielsen & R. Friedman. Dordrecht: Kluwer, 9–36, 2003.

Hull, Clark L. *Principles of Behavior: An Introduction to Behavior Theory*. New York: Appleton-Century, 1943.

Husserl, Edmund. *Aufsätze und Rezensionen (1890–1910)*. Husserliana XXII. Den Haag: Martinus Nijhoff, 1979.

Husserl, Edmund. *Aufsätze und Vorträge (1911–1921)*. Husserliana XXV. Den Haag: Martinus Nijhoff, 1987.

Husserl, Edmund. *Briefwechsel*. Husserliana Dokumente III/1–10. Dordrecht: Kluwer Academic Publishers, 1994.

Husserl, Edmund. *Cartesian Meditations: An Introduction to Phenomenology*. Translated by Dorion Cairns. Dordrecht, Boston, Lancaster: Martinus Nijhoff, 1960.

Husserl, Edmund. *Cartesianische Meditationen und Pariser Vorträge*. Husserliana I. Den Haag: Martinus Nijhoff, 1950.

Husserl, Edmund. *The Crisis of European Sciences and Transcendental Phenomenology*. Translated by David Carr. Evanston: Northwestern University Press, 1970.

Husserl, Edmund. *Erfahrung und Urteil: Untersuchungen zur Genealogie der Logik*. Edited by Ludwig Landgrebe. Hamburg: Felix Meiner Verlag, 1999 [1939]. An English translation *Experience and Judgment: Investigations in the Genealogy of Logic*. Translated by James S. Churchill and Karl Ameriks. Evanston: Northwestern University Press, 1973.

Husserl, Edmund. *Erste Philosophie (1923/24). Erster Teil. Kritische Ideengeschichte*. Husserliana VII. Den Haag: Martinus Nijhoff, 1956.

Husserl, Edmund. *Erste Philosophie (1923/24). Zweiter Teil. Theorie der phänomenologischen Reduktion*. Husserliana VIII. Den Haag: Martinus Nijhoff, 1959.

Husserl, Edmund. *Formale und Transzendentale Logik*. Husserliana XVII. Den Haag: Martinus Nijhoff, 1974.

Husserl, Edmund. *Die Idee der Phänomenologie. Fünf Vorlesungen*. Husserliana II. Den Haag: Martinus Nijhoff, 1950.

Husserl, Edmund. *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Drittes Buch: Die Phänomenologie und die Fundamente der Wissenschaften*. Husserliana V. Den Haag: Martinus Nijhoff, 1952.

Husserl, Edmund. *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie Erstes Buch: Allgemeine Einführung in die reinen Phänomenologie*. Husserliana III/1–2. Den Haag: Martinus Nijhoff, 1976.

Husserl, Edmund. *Ideen zu einer reinen Phänomenologie und phänomenologishen Philosophie. Zweites Buch: Phänomenologische Untersuchungen zur Konstitution*. Husserliana IV. Den Haag: Martinus Nijhoff, 1952. An English translation *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, Second Book: Studies in the Phenomenology of Constitution*. Translated by R. Rojcewicz and A. Schuwer. Dordrecht, Boston, London: Kluwer Academic Publishers, 1989.

Husserl, Edmund. *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie. Eine Einleitung in die phänomenologische Philosophie*. Husserliana VI. Den Haag: Martinus Nijhoff, 1954.

Husserl, Edmund. *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie. Ergänzungsband. Texte aus dem Nachlass 1934–1937*. Husserliana XXIX. Dordrecht: Kluwer Academic Publishers, 1993.

Husserl, Edmund. *Logical Investigations. Volume I*. Translated by J.N. Findlay. London and New York: Routledge, 1982.

Husserl, Edmund. *Logische Untersuchungen. Erster Band. Prolegomena zur reinen Logik*. Husserliana XVIII. Den Haag: Martinus Nijhoff, 1975. An English translation *Logical Investigations. Volume I*. Translated by J.N. Findlay. London and New York: Routledge, 1982.

Husserl, Edmund. *Logische Untersuchungen. Zweiter Band. Untersuchungen zur Phänomenologie und Theorie der Erkenntnis*. Husserliana XIX/1–2. Den Haag: Martinus Nijhoff, 1984.

Husserl, Edmund. *Natur und Geist. Vorlesungen Sommersemester 1927*. Husserliana XXXII. Dordrecht: Kluwer Academic Publishers, 2001.

Husserl, Edmund. "Phenomenology." Translated by C.V. Salmon. *The Encyclopaedia Britannica*, 14th ed., vol. 17: 699–702, 1929. Also published in *Psychological and Transcendental Phenomenology and the Confrontation with Heidegger (1927–1931). Collected Works: Volume 6*. Edited and translated by Thomas Sheehan and Richard E. Palmer. Dordrecht: Kluwer, 1997.

Husserl, Edmund. *Phänomenologische Psychologie*. Husserliana IX. Den Haag: Martinus Nijhoff, 1962.

Huxley, T.H. *Science and Culture*. New York: Appleton, 1884.

Jackson, Frank. "What Mary Didn't Know." In *The Nature of Mind*. Edited by David Rosenthal. New York: Oxford University Press, 291–95, 1991.

Jackson, Frank. *From Metaphysics to Ethics*. Oxford: Oxford University Press, 1998.

James, William. *Essays in Radical Empiricism*. Edited by Ralph Barton Perry. Lincoln and London: University of Nebraska Press, 1996 [1912].

James, William. *The Meaning of Truth: A Sequel to Pragmatism*. In *The Works of William James*, 19 vols. Edited by H. Burkhardt, Fredson Bowers and Ignas K. Skrupskelis. Cambridge, MA and London: Harvard University Press, 1978 [1909].

James, William. *A Pluralistic Universe* [1909]. In *The Works of William James*, 19 vols. Edited by H. Burkhardt, Fredson Bowers and Ignas K. Skrupskelis. Cambridge, MA and London: Harvard University Press, 1978.

James, William. *Pragmatism: A New Name for Some Old Ways of Thinking*. In *The Works of William James*, 19 vols. Edited by H. Burkhardt, Fredson Bowers and Ignas K. Skrupskelis. Cambridge, MA and London: Harvard University Press, 1978.

James, William. *The Varieties of Religious Experience* (1902 [1985]). In *The Works of William James*, 19 vols. Edited by H. Burkhardt, Fredson Bowers and Ignas K. Skrupskelis. Cambridge, MA and London: Harvard University Press, 1978.

Janduno, John, see John of Janduno.

Javelli, Chrysostomus, see Chrysostomus Javelli.

Jerry, Fodor. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge: The MIT Press, 1987.

Jodocus Trutfetter. *Summa in Totam Physicen*. Erfurt: Mathias Maler, 1514.

Jodocus Trutfetter. *Veteris artis: id est Porpyrii universalium: et praedicamentorum Aristotelis: Perihermeneiasque expositio tam brevis quam utilis una cum dubiolorum ex iis extractorum resolutione*. Erfurt (?), ca. 1500.

Johannes Carnificis de Lutrea. *Exercitium librorum de anima*. Erfurt: Paulus Wider de Hornbach, 1482.

Johannes Peyligk. *Philosophiae naturalis compendium*. Leipzig: Melchior Lotter, 1499.

John Buridan. *Quaestiones in tres libros De anima Aristotelis* (Tertia lectura). In *John Buridan on the Soul and Sensation. An Edition of Book II of His Commentary on Aristotle's Book on the Soul with an Introduction and a Translation of Question 18 on Sensible Species*. Edited by translated by Peter Gordon Sobol. Dissertation at Indiana University, 1984.

John Buridan. *La traité de l'âme [De prima lectura]*. Edited by Benoît Patar. Louvain-la-Neuve: Éditions de l'Institut supérieur de philosophie, 1991.

John Duns Scotus. *Opera Omnia*, Vol. 7. Edited by Luke Wadding. Lyon, 1639.

John Eck. *Aristotelis Stagyritae philosophi De anima libri tres adiectis Eckii Commentariis*. Augsburg, 1520.

John of Janduno. *Super libros Aristotelis De anima quaestiones*. Venice, 1587.

Juti, Riku. *Pioneering Externalism: Thomas Reid and Hume's Problem*. Reports from the Department of Philosophy, Vol. 1. University of Helsinki, 1993.

Kant, Immanuel. *Anthropology from a Pragmatic Point of View*. Translated, with an introduction and notes by Mary Gregor. Den Hague: Martinus Nijhoff, 1974.

Kant, Immanuel. *Critique of the Power of Judgment*. Edited and translated by Paul Guyer and Eric Matthews. Cambridge: Cambridge University Press, 2000.

Kant, Immanuel. *Critique of Practical Reason*. Translated and edited by Mary Gregor. Cambridge: Cambridge University Press, 2002.

Kant, Immanuel. *Critique of Pure Reason*. Edited and translated by Paul Gruyer and Allen W. Wood. Cambridge: Cambridge University Press, 1998.

Kant, Immanuel. *Kant's gesammelte Schriften*. Berlin: Königlichen Preussischen Akademie der Wissenschaften, 1910.

Kärkkäinen, Pekka. "Nominalist Psychology and the Limits of Canon Law in Late Medieval Erfurt." In *Lutheran Reformation and the Law*. Edited by Virpi Mäkinen. Leiden: Brill, 93–110, 2006.

Kern, Iso. *Husserl und Kant: Eine Untersuchung über Husserls Verhältnis zu Kant und zum Neukantianismus*. Den Haag: Martinus Nijhoff, 1964.

Kessler, Eckhard. "Intellective Soul." In *The Cambridge History of Renaissance Philosophy*. Edited by C. Schmitt et al. Cambridge: Cambridge University Press, 486–534, 1988.

Kim, Jaegwon. *Philosophy of Mind*. Dimension of Philosophy Series. Boulder: Westview, 1996.

Kitcher, Patricia. "Kant's Real Self." In *Self and Nature in Kant's Philosophy*, Edited by Allen W. Wood. Ithaca and London: Cornell University Press, 113–47, 1984.

Kitcher, Patricia. *Kant's Transcendental Psychology*. Oxford: Oxford University Press, 126–27, 1990.

Kleineidam, Erich. *Universitas Studii Erffordensis I*. Leipzig: St. Benno-Verlag, 1964.

Kleineidam, Erich. *Universitas Studii Erffordensis II*. Erfurt: Benno, 1992.

Knuuttila, Simo. *Emotions in Ancient and Medieval Philosophy*. Oxford: Clarendon Press, 2004.

Koffka, Kurt. *Principles of Gestalt Psychology*. New York: Harcourt, Brace, 1935.

Kripke, Saul. *Wittgenstein on Rules and Private Language*. Oxford: Basil Blackwell, 1982.

Kristeller, P.O. "The Myth of Renaissance Atheism and the French Tradition of Free Thought." *Journal of the History of Philosophy* 6: 233–43, 1968.

Kristeller, P.O. "Paduan Averroism and Alexanderism in the Light of Recient Studies." In *Aristotelismo padovano e filosofia aristotelica, Atti del XII Congresso Internazionale di Filosofia, IX*, Florence, 147–55, 1960. Also published in *Renaissance Thought* II. New York: Harper & Row, 1965.

Kuksewicz, Zdzislaw. "Paul de Venise et sa théorie de l'âme." In *Aristotelismo veneto e scienza moderna* I. Edited by L. Olivieri. Padova: Antenore, 1983.

Kusch, Martin. *Psychologism: A Case Study in the Sociology of Philosophical Knowledge*. London: Routledge, 1995.

Lachelier, Jules. "Psychologie et métaphysique." In *Œuvres de Lachelier*. Paris: Alcan, 1896.

Lagerlund, Henrik. "John Buridan and the Problems of Dualism in the Early Fourteenth Century." *Journal of the History of Philosophy* 42: 369–87, 2004.

Lagerlund, Henrik. "Representations, Concepts and Words – Peter of Ailly on Semantics and Psychology." *Proceedings of the Society for Medieval Logic and Metaphysics* 3: 15–36, 2003.

Laudan, Larry. *Science and Hypotheses*. Dordrecht: Reidel, 86–110, 1981.

Lavater, Johann Caspar. *Essays on Physiognomy; For the Promotion of the Knowledge and the Love of Mankind*. Translated by Thomas Holcroft. London: G. G. J. & J. Robinson, 1789.

Le Grand, Antoine. *An Entire Body of Philosophy, According to the Principles of the Famous Renate Des Cartes*. Translated by Richard Blome. London, 1694; repr., New York: Johnson, 1972.

Le Grand, Antoine. *Institutio philosophiae secundum principia de Renati Descartes*. London, 1678.

Leahey, Thomas H. "David Hartley." In *The Cambridge Dictionary of Philosophy*. Edited by R. Audi. Cambridge: Cambridge University Press, 1995.

Locke, John. *Essay Concerning Human Understanding*. Edited by P.H. Nidditch. Oxford: Clarendon Press, 1975.

Lohr, Charles H. "Renaissance Latin Aristotle Commentaries." *Studies in the Renaissance* 21: 1974, and *Renaissance Quarterly* 28–33: 1975–1980, and 35: 1982.

Longuenesse, B. *Kant and the Capacity to Judge: Sensibility and Discursivity in the Transcendental Analytic of the Critique of Pure Reason*. Princeton: Princeton University Press, 1998.

Lundberg, Mabel. *Jesuitische Anthropologie und Erziehungslehre in der Frühzeit des Ordens (ca. 1540 – ca. 1650)*. Studia Doctrinae Christianae Upsaliensia 6. Uppsala: Almqvist & Wiksell, 1966.

Lyons, William. *Disappearance of Introspection*. Cambridge: The MIT Press, 1986.

Mach, Ernst. *The Analysis of Sensations and the Relation of the Physical to the Psychical*. Translated by C.M. Williams and S. Waterlow. New York: Dover, 1959.

Mach, Ernst. *Beiträge zur Analyse der Empfindungen*. Jena, 1886.

Madison, Gary. *The Phenomenology of Merleau-Ponty*. Athens: Ohio University Press, 1981 [1973].

Mahoney, Edward P. "Agostino Nifo (ca. 1470–1538) on the *scientia de anima* as a 'mathematical' or 'middle' science." In *Knowledge and the Sciences in Medieval Philosophy III*. Edited by R. Työrinoja et al. Helsinki: Luther-Agricola Society, 629–36, 1990.

Mahoney, Edward P. "Albert the Great and the Studio Patavino in the Late Fifteenth and Early Sixteenth Centuries." In *Albertus Magnus and the Sciences*. Edited by J.A. Weisheipl. Toronto: Pontifical Institute of Mediaeval Studies, 537–66, 1980.

Mahoney, Edward P. "Antonio Trombetta and Agostino Nifo on Averroes and Intelligible Species: A Philosophical Dispute at the University of Padua." In *Storia e Cultura al Santo di Padova fra il XIII e il XX secolo*. Edited by A. Poppi. Venezia, 289–301, 1976.

Mahoney, Edward P. "Nicoletto Vernia and Agustino Nifo on Alexander of Aphrodisias: An Unnoted Dispute." *Rivista Critica di Storia della Filosofia* 23: 268–96, 1968.

Mahoney, Edward P. "Nicoletto Vernia on the Soul and Immortality." In *Philosophy and Humanism: Renaissance Essays in Honor of Paul Oskar Kristeller*. Edited by E.P. Mahoney. New York and Leiden, 144–63, 1976.

Mahoney, Edward P. "Saint Thomas and the School of Padua at the End of the Fifteenth Century." *Proceeding of the American Catholic Philosophical Association: Thomas and Bonaventure, A Septicentenary Commemoration* 48: 277–85, 1974.

Maier, Anneliese. "Das Problem der 'species sensibiles in medio' und die neue Naturphilosophie des 14. Jahrhunderts." *Ausgehendes Mittelalter II*. 419–51, 1967.

Malagola, Carlo (ed.). *Statuti delle universitá e dei collegi dello Studio bolognese*. Bologna, 1966.

Malpas, Jeff. *From Kant to Davidson: Philosophy and the Idea of the Transcendental*. London and New York: Routledge, 2003.

Marshall, Peter. "Parisian Psychology in the Mid-Fourteenth Century." In *Archives d'histoire doctrinale et littéraire du moyen âge* 50: 101–93, 1984.

Martinelli, Riccardo. *Misurare l'anima, filosofia e psicofisica da Kant a Carnap*. Macerata: Quodlibet, 1999.

Matsen, Herbert. "Alessandro Achillini (1463–1512) and 'Ockhamism' at Bologna (1490–1500)." *Journal of the History of Philosophy* 13: 437–51, 1975.

McDowell, John. "Having the World in View." *The Journal of Philosophy* 95: 431–91, 1998.

McDowell, John. "Two Sorts of Naturalism." In *Mind, Value, and Reality*. Edited by John McDowell. Cambridge: Harvard University Press, 167–97, 1998.

McGinn, Marie. *Wittgenstein and the Philosophical Investigations*. London: Routledge, 1997.

Merleau-Ponty, Maurice. "Le métaphysique dans l'homme." In *Sens et non-sens*. Paris: Gallimard, 104–06, 1996 [1966]. An English translation "The Metaphysical in Man." In *Sense and Non-Sense*. Translated by Hubert L. Dreyfus and Patricia Allen Dreyfus. Evanston. Illinois: Northwestern University Press, 85–6, 1964.

Merleau-Ponty, Maurice. *Phénoménologie de la perception*. Paris: Gallimard, 1945. An English translation *Phenomenology of Perception*. Translated by Colin Smith. London and New York: Routledge, 1995.

Merleau-Ponty, Maurice. *L'union de l'âme et du corps chez Malebranche Biran et Bergson*. Paris: Gallimard, 1968.

Mikkeli, Heikki. *An Aristotelian Response to Renaissance Humanism*. Helsinki: Suomen Historiallinen Seura, 1992.

Mohanty, Jitendranath. *The Possibility of Transcendental Philosophy*. Dordrecht: Martinus Nijhoff, 1985.

Mohanty, Jitendranath. *Transcendental Phenomenology: An Analytic Account*. Oxford and Cambridge, MA: Blackwell, 1989.

Monfasani, John. "Aristotelians, Platonists, and the Missing Ockhamists: Philosophical Liberty in Pre-Reformation Italy." *Renaissance Quarterly* 46: 247–76, 1993.

Moore, Francis Charles Timothy. *Bergson: Thinking Backwards*. Cambridge: Cambridge University Press, 1996.

Moore, Francis Charles Timothy. *The Psychology of Maine de Biran*. Oxford: Clarendon Press, 1970.

Moran, Dermot. "Making Sense: Husserl's Phenomenology as Transcendental Idealism." In *From Kant to Davidson: Philosophy and the Idea of the Transcendental*. Edited by Jeff Malpas. London: Routledge, 48–74, 2002.

Moran, Richard. *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton: Princeton University Press, 2001.

Movia, Giancarlo. "Note sul primo capitolo del 'De anima' di Aristotele." In *Scritti in onore di Carlo Giacon*. Padua: Antenore, 1972.

Müller, Johannes. *Handbuch der Physiologie des Menschen*. Koblenz: Hölscher, 1840.

Münch, Dieter. *Intention und Zeichen: Untersuchungen zu Franz Brentano und zu Edmund Husserls Frühwerk*. Frankfurt: Suhrkamp, 1993.

Mulligan, Kevin and Smith, Barry. "Mach and Ehrenfels: The Foundations of Gestalt theory." In *The Foundations of Gestalt Theory*. Edited by Barry Smith München. Wien: Philosophia Verlag, 124–57, 1988.

Nagel, Thomas. "What Is It Like To Be a Bat?" *The Philosophical Review* 4: 435–50, 1974.

Nardi, Bruno. *Saggi sull'aristotelismo padovano del secolo XIV al XVI*. Florence, 1961.

Nicoletto Vernia. *Contra perversam Averrois opinionem de unitae intellectus et de animae felicitate questiones divinae*. Venice, 1505.

Nifo, Agostino, see Agostino Nifo.

O' Saughnessy, Brian. *The Will*. Cambridge: Cambridge University Press, 1980.

Park, Katherine. "Albert's Influence on Late Medieval Psychology." In *Albertus Magnus and the Sciences: Commemorative Essays 1980*. Edited by James A. Weisheipl. Toronto: Pontifical Institute of Medieval Studies, 510–22, 1980.

Park, Katharine. "The Organic Soul." In *The Cambridge History of Renaissance Philosophy*. Edited by C. Schmitt et al. Cambridge: Cambridge University Press, 464–84, 1988.

Pasnau, Robert. "Cognition." In *The Cambridge Companion to Duns Scotus*. Edited by T. Williams. Cambridge: Cambridge University Press, 285–311, 2003.

Pawelski, James O. "William James's Divided Self and the Process of Its Unification: A Reply to Richard Gale." *Transactions of the Charles S. Peirce Society* 39: 645–56, 2003.

Peirce, Charles S. *The Collected Papers of Charles Sanders Peirce*, 8 vols. Edited by Charles Hartshorne and Paul Weiss (vols. 16) and Arthur W. Burks (vols. 7–8). Cambridge, MA: Harvard University Press, 1931–1958.

Peirce, Charles S. *The Essential Peirce*. The Peirce Edition Project, 2 Vols. Bloomington: Indiana University Press, 1992–1998.

Perry, Ralph Barton. *The Thought and Character of William James: Briefer Version*. New York: Harper & Row, 1964 [1935].

Peters, F.E. *Greek Philosophical Terms: A Historical Lexicon*. New York: New York University Press, 1967.

Peyligk, Johannes, see Johannes Peyligk.

Philaltheus, Lucillus (Lucillo Maggi Filalteo). *In libros tres Aristotelis De anima Commentaria absolutissima.* Turin, 1579.

Piccolomini, Franciscus. *Commentarium in tres libros Aristotelis De anima.* Frankfurt, 1602.

Pietro Pomponazzi. "On the Immortality of the Soul." Translated by W.H. Hay. In *The Renaissance Philosophy of Man*. Edited by E. Cassirer, P.O. Kristeller and J.H. Randall Jr. Chicago: The University of Chicago Press, 1998.

Pihlström, Sami. "Mortality, Individuality, and Pluralism: William James's Democratic Religion." *American Journal of Theology and Philosophy* 26: 96–120, 2005.

Pihlström, Sami. *Naturalizing the Transcendental: A Pragmatic View*. Amherst, NY: Prometheus/Humanity Books, 2003.

Pihlström, Sami. "On the Reality of Evil: A Jamesian Investigation." *Streams of William James* 4: 12–21, 2002.

Pihlström, Sami. "Peirce's Place in the Pragmatist Tradition." in *The Cambridge Companion to Peirce*. Edited by Cheryl Misak. Cambridge: Cambridge University Press, 27–57, 2004.

Pihlström, Sami. *Pragmatic Moral Realism: A Transcendental Defense*. New York and Amsterdam: Rodopi, 2005.

Pihlström, Sami. "Pragmatism and American Personalism: On the Possibility of Perspectival Metaphysics." *Iyyun* 53: 287–324, 2004.

Pihlström, Sami. *Pragmatism and Philosophical Anthropology: Understanding Our Human Life in a Human World*. New York: Peter Lang, 1998.

Pihlström, Sami. *Solipsism: History, Critique, and Relevance*. Tampere: Tampere University Press, 2004.

Pihlström, Sami. *Structuring the World: The Issue of Realism and the Nature of Ontological Problems in Classical and Contemporary Pragmatism*. Acta Philosophica Fennica 59. Helsinki: The Philosophical Society of Finland, 1996.

Pihlström, Sami. "William James on Death, Mortality, and Immortality." *Transactions of the Charles S. Peirce Society* 38: 605–28, 2002.

Pike, Martin L. *Pietro Pomponazzi: Radical Philosopher of the Renaissance*. Padova: Antenore, 1986.

Pink, Thomas. "Action, Will and Law in Late Scholasticism." In *Ethics on the Threshold of Modernity*. Edited by J. Kraye and R. Saarinen. Dordrecht: Springer, 91–111, 2004.

Pink, Thomas. *The Ethics of Action*. Vol. 1: *Self-Determination* and Vol. 2: *Normativity*. Oxford: Oxford University Press, forthcoming.

Pink, Thomas. "Moral Obligation." In *Modern Moral Philosophy*. Edited by A. O'Hear. Cambridge: Cambridge University Press, 80–100, 2004.

Pink, Thomas. "Suarez, Hobbes and the Scholastic Tradition in Action Theory." In *The Will and Human Action: From Antiquity to the Present Day*. Edited by T. Pink and M. Stone. London: Routledge, 127–53, 2003.

Pippin, Robert B. "Kant and the Spontaneity of Mind." *Canadian Journal of Philosophy* 17: 449–75, 1987.

Pomponazzi, Pietro, see Pietro Pomponazzi.

Poppi, Antonio. *Introduzione all'aristotelismo padovano*. Padova: Antenore, 1970.

Poppi, Antonio (ed.). *Scienza e filosofia all'Università di Padova nel Quattrocento*. Trieste: Lint, 1983.

Price, Richard. *A Review of the Principal Questions in Morals*, 3rd ed. London: T. Cadell in the Strand, 1787.

Proudfoot, Wayne (ed.). *William James and a Science of Religions: Reexperiencing The Varieties of Religious Experience*. New York: Columbia University Press, 2004.

Putnam, Hilary. *The Collapse of the Fact/Value Dichotomy and Other Essays*. Cambridge, MA and London: Harvard University Press, 2002.

Putnam, Hilary. *Ethics without Ontology*. Cambridge, MA and London: Harvard University Press, 2004.

Putnam, Hilary. "Levinas and Judaism." In *The Cambridge Companion to Levinas*. Edited by Simon Critchley and Robert Bernasconi. Cambridge: Cambridge University Press, 33–62, 2002.

Putnam, Hilary. "Philosophy as a Reconstructive Activity: William James on Moral Philosophy." In *The Pragmatic Turn in Philosophy: Contemporary Engagements between Analytic and Continental Thought*. Edited by William Eggington and Mike Sandbothe. Albany: SUNY Press, 31–46, 2004.

Putnam, Hilary. "Sense, Nonsense, and the Senses: An Inquiry into the Powers of the Human Mind." *The Journal of Philosophy* 91: 445–517, 1994.

Putnam, Hilary. *The Threefold Cord: Mind, Body, and World*. New York: Columbia University Press, 1999.

Putnam, Hilary. *Renewing Philosophy*. Cambridge: Harvard University Press, 1992.

Questio preambularis in disputacionem de quolibet celebratam Erffordie A. D. 1497. Edited by Hans-Ulrich Wöhler. *Bochumer philosophisches Jahrbuch für Antike und Mittelalter* 6: 138–95, 2001.

Quine, W.V.O. *The Roots of Reference*. LaSalle, Ill.: Open Court, 1973.

Quine, W.V.O. *Word and Object*. Cambridge: The MIT Press, 1960.

Radner, Daisie. "The Function of the Passions." In *Passion and Virtue in Descartes*. Edited by B. Williston and A. Gombay. New York: Humanity Books, 175–87, 2003.

Ragnisco, Pietro. *Documenti inediti e rari intorno alla vita ed agli scritti di Nicoletto Vernia e di Elia del Medigo*. Padua, 1891.

Randall, J.H. *The Making of the Modern Mind*. New York: Colombia University Press, 1976.

Randall, J.H. "Paduan Aristotelianism Reconsidered." In *Philosophy and Humanism: Renaissance Essays in Honor of Paul Oskar Kristeller*. Edited by E.P. Mahoney. New York and Leiden, 275–82, 1976.

Randall, J.H. *The School of Padua and the Emergence of Modern Science*. Padua: Editrice Antenore, 1961.

Regis, Pierre Sylvain. *Système de philosophie: contenant la logique, metaphysique, physique & morale*, 7 vols. Lyon: Anisson, Posuel & Rigaud, 1691.

Reid, Thomas. *Philosophical Works*. 2 Vols. Edited by Sir William Hamilton. Edinburgh: James Thin, 1895.

Reuter, Martina. "Catharine Macaulay and Mary Wollstonecraft on the Will." In *Virtue, Liberty, and Toleration: Political Ideas of European Women, 1400–1800*. Edited by Jacqueline Broad and Karen Green. Dordrecht: Springer, 149–69, 2007.

Reuter, Martina. "Mary Wollstonecraft on Love and Friendship." In *Philosophical Aspects on Emotions*. Edited by Åsa Carlson. Stockholm: Thales, 119–39, 2005.

Rey, Georges. *Contemporary Philosophy of Mind*. Oxford: Blackwell, 1997.

Ribot, Theodule. "Philosophy in France." *Mind* 2: 366–82, 1877.

Rivers, Christopher. *Face Value: Physiognomical Thought and the Legible Body in Marivaux, Lavater, Balzac, Gautier, and Zola*. Madison: The University of Wisconsin Press, 1994.

Rosenstock, Gershon George. *F. A. Trendelenburg: Forerunner to John Dewey*. Carbondale: Southern Illinois University Press, 1964.

Rosenthal, David. "Introspection and Self-Interpretation." *Philosophical Topics* 28: 201–33, 2000.

Rosenthal, Sandra B. *Speculative Pragmatism*. Amherst: The University of Massachusetts Press, 1986.

Royce, Josiah. *The Problem of Christianity*. Washington, D.C.: The Catholic University Press of America, 2001; reprint from the University of Chicago Press edition, 1968 [1913].

Russell, Bertrand. *Analysis of Mind*. London: Allen & Unwin, 1921.

Russell, Bertrand. *Mysticism and Logic*. London: Allen and Unwin, 1917.

Russell, John E. "Solipsism: The Logical Issue of Radical Empiricism." *The Philosophical Review* 15: 606–13, 1906.

Sachs-Hombach, Klaus. *Philosophische Psychologie im 19. Jahrhundert: Ihre Entstehung und Problemgeschichte*. Freiburg-München: Alber, 1993.

Sapiro, Virginia. *A Vindication of Political Virtue: The Political Theory of Mary Wollstonecraft*. Chicago: University of Chicago Press, 1992.

Sartre, Jean-Paul. *Being and Nothingness: A Phenomenological Essay on Ontology*. Translated by Hazel E. Barnes. New York: Philosophical Library, 1966 [1943].

Scanlon, T.M. *What We Owe to Each Other*. Cambridge: Harvard University Press, 1998.

Schaper, Eva. "Taste, Sublimity, and Genius: The Aesthetics of Nature and Art." In *The Cambridge Companion to Kant*. Edited by Paul Guyer. Cambridge: Cambridge University Press, 367–93, 1992.

Scheerer, Eckart. "Psychologie." In *Historisches Wörterbuch der Philosophie* vol. 7. Edited by Joachim Ritter. Basel: Schwabe, 1971, cols. 1599–1653.

Schmitt, C.B. *Aristotle and the Renaissance*. Cambridge, Mass.: Harvard University Press, 1983.

Scipione Dindi Dall'Orologio, Francesco. *Dissertazioni sopra l'istoria ecclesiastica di Padova: Dissertazione nona*. Padua, 1817.

Scotus, John Duns, see John Duns Scotus.

Seigfried, Charlene Haddock. *William James's Radical Reconstruction of Philosophy*. Albany: SUNY Press, 1990.

Serck-Hanssen, Camilla. "Apperception and Deduction in the *Duisburg'sche Nachlass*." In *Kant und die Berliner Aufklärung*. Akten des 9. Internationalen Kant-Kongresses, Vol. 2. Berlin: de Gruyter, 59–69, 2001.

Serck-Hanssen, Camilla. "Kant's Critical Debut." In *From Kant to Davidson: Philosophy and the Idea of the Transcendental*. Edited by Jeff Malpas. London: Routledge, 7–22, 2003.

Shapiro, Alan E. "Newton's Optics and Atomism." In *The Cambridge Companion to Newton*. Edited by I. Bernard Cohen and George E. Smith. Cambridge: Cambridge University Press, 227–55, 2002.

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

Shoemaker, Sydney. "Introspection and Phenomenal Character." *Philosophical Topics* 28: 247–73, 2000.

Shook, John. *Pragmatism: An Annotated Bibliography 1898–1940*. Amsterdam and Atlanta: Rodopi, 1998.

Shortland, Michael. "Skin Deep: Barthes, Lavater and the Legible Body." *Economy and Society* 14: 273–312, 1985.

Skinner, B.F. "Behaviorism at Fifty." *Science* 140: 951–58, 1963.

Skinner, B.F. "The Operational Analysis of Behavioral Terms." *Psychological Review* 52: 270–77, 1945.

Smith, Barry. *Austrian Philosophy: The Legacy of Franz Brentano*. Chicago, Illinois: Open Court Publishing Company, 1994.

Smith, Barry. *The Foundations of Gestalt Theory*. München, Wien: Philosophia Verlag, 1988.

Smith, Roger. "The History of Psychological Categories." *Studies in History and Philosophy of Biological and Biomedical Sciences* 36: 55–94, 2005.

Sokolowski, Robert. *The Formation of Husserl's Concept of Consciousness*. The Hague: Martinus Nijhoff, 1964.

Sokolowski, Robert. "Logic of Parts and Wholes." *Philosophy and Phenomenological Research* 28: 537–53, 1968.

Sorabji, Richard. "From Aristotle to Brentano: The Development of the Concept of Intentionality." In *Oxford Studies in Ancient Philosophy. Supplementary Volume: Aristotle and the Later Tradition*. Edited by Henry Blumenthal and Henry Robinson. Oxford: Clarendon Press, 1991.

Spade, P.V. "Insolubilia." In *The Cambridge History of Later Medieval Philosophy*. Edited by N. Kretzmann, A. Kenny and J. Pinborg. Cambridge: Cambridge University Press, 1982.

Spinoza, Baruch. *Opera*, 5 vols. Heidelberg: Winter, 1925–1938.

Spinoza, Baruch. *Verhandeling over de verbetering van het verstand*. Translated by Theo Verbeek. Groningen: Historische Uitgeverij, 2002.

Sprague, Elmer. *Persons and their Minds: A Philosophical Inquiry*. Boulder, CO: Westview Press, 1999.

Sprigge, T.L.S. *James and Bradley: American Truth and British Reality*. Chicago and La Salle, IL: Open Court, 1993.

Spruit, Leen. *Species Intelligibilis: From Perception to Knowledge*. Vol. 2. Leiden: Brill, 1995.

Stallmach, Josef. *Dynamis und Energeia*. Meisenheim am Glan: Anton Hain, 1959.

Stawarska, Beata. "Anonymity and Sociality: The Convergence of Psychological and Philosophical Currents in Merleau-Ponty's Ontological Theory of Intersubjectivity." *CHIASMI International* 5: 295–309, 2004.

Steinbock, Anthony. *Home and Beyond: Generative Phenomenology after Husserl*. Evanston, Illinois: Northwestern University Press, 1973.

Stoecker, Ralf (ed.) *Reflecting Davidson*. Berlin: de Gruyter, 1993.

Stoutland, Frederick. "Analytic Philosophy and Metaphysics." In *Wittgenstein and the Method of Philosophy*. Acta Philosophica Fennica, Vol. 80. Edited by Sami Pihlström. Helsinki, 2006.

Stoutland, Frederick. "Davidson on Intentional Behavior." In *Actions and Events*. Edited by E. Lepore and B. McLaughlin. Oxford: Basil Blackwell, 44–59, 1985.

Stoutland, Frederick. "Intentionalists and Davidson on Rational Explanation." In *Actions, Norms and Values*. Edited by G. Meggle. Berlin: Walter de Gruyter, 191–208, 1991.

Strawson, Galen. *Mental Reality*. Cambridge: The MIT Press, 1994.

Strawson, Peter F. *Individuals*. London: Methuen, 1959.

Sturm, Thomas. "Kant on Psychology: How Not to Investigate the Human Mind." In *Kant and the Sciences*. Edited by Eric Watkins. New York: Oxford University Press, 163–84, 2001.

Suárez, Francisco. *Commentaria una cum quaestionibus in libros Aristotelis De anima*. Ediciones críticas de obras filosóficas 1. Edited by Salvador Castellote Cubells. Madrid: Labor, 1978.

Suárez, Francisco. *Disputationes metaphysicae*. Paris, 1866; repr., Hildesheim: Georg Olms, 1965.

Suárez, Francisco. *Opera Omnia*. Vol. 5: *De Legibus ac Legislatore Deo*. Edited by Charles Berton. Paris: Vives, 1856.

Sutton, T.J. "The Scottish Kant?" In *The Philosophy of Thomas Reid*. Edited by M. Dalgarno and E. Matthews. Dordrecht: Kluwer Academic Publishers, 159–92, 1989.

Tachau, Catherine H. "Some Aspects of Intentional Existence at Paris 1250–1320." In *Medieval Analyses in Language and Cognition. Acts of the Symposium The Copenhagen School of Medieval Philosophy January 10–13, 1996*. Edited by S. Ebbesen and R.L. Friedman. Historisk-filosofiske Meddelelser 77. Copenhagen: The Royal Danish Academy of Sciences and Letters, 331–53, 1999.

Taylor, Charles. *Varieties of Religion Today: William James Revisited*. Cambridge, MA and London: Harvard University Press, 2002.

Tewes, Götz-Rüdiger. "Die Erfurter Nominalisten und ihre thomistischen Widersacher in Köln, Leipzig und Wittenberg. Ein Beitrag zum deutschen Humanismus am Vorabend der Reformation." In *Die Bibliotheca Amploniana*. Edited by Andreas Speer. Berlin: Walter de Gruyter, 447–88, 1995.

Thomas Aquinas. *A Commentary on Aristotle's De anima*. Translated by Robert Pasnau. New Haven: Yale University Press, 1999.

Thomas Aquinas. *In quattuor libros Sententiarum* II.In *Opera Omnia* 1. Edited by Roberto Busa. Stuttgart: Friedrich Frommann Verlag Günther Holzboog, 1980.

Thomas Aquinas. *Summa Theologiae*. Turin: Marietti, 1950.

Thomas Cajetanus (Jacopo Vio). *Commentaria in De anima Aristotelis*, Vol. I. Edited by P. I. Coquelle. Rome, 1938.

Thomas Cajetanus (Jacopo Vio). *Commentaria in libros Aristotelis De anima III*. Edited by G. Picard and G. Pelland. Paris: Desclée de Brouwer, 1965.

Tipton, Ian. " 'Ideas' and 'Objects': Locke on Perceiving 'Things'." In *The Empiricists: Critical Essays on Locke, Berkeley, and Hume*. Edited by Margaret Atherton. Oxford: Oxford University Press, 97–110, 1999.

Tipton, Ian. "Blasius of Parma". In *Routledge Encyclopedia of Philosophy*. Edited by E. Craig. London: Routledge, 1998.

Toadvine, Ted. "Merleau-Ponty's Reading of Husserl." In *Merleau-Ponty's Reading of Husserl*. Edited by Ted Toadvine and Lester Embree. Dordrecht, Boston, London: Kluwer, 227–86, 2002.

Toadvine, Ted. "Phenomenological Method in Merleau-Ponty's Critique of Gurwitsch." *Husserl Studies* 17: 195–205, 2001.

Todd, Janet. *Mary Wollstonecraft: A Revolutionary Life*. London: Weidenfeld & Nicolson, 2000.

Todd, Janet. "Prefatory Note." In Wollstonecraft, *The Works of Mary Wollstonecraft*, Vol. 7. Edited by Janet Todd and Marilyn Butler. London: William Pickering, 1989.

Toletus, Franciscus. *Commentaria una cum quaestionibus in tres libros Aristotelis De anima*. Cologne, 1615.

Trendelenburg, Friedrich Adolf. *Aristoteles de anima libri tres: Ad interpretum graecorum auctoritatem et codicum fidem recognivit, commentariis illustravit F. A. Trendelenburg*. Jena: Walz, 1833; repr., Berlin: Weber, 1877; Graz: Akademische Druck- und Verlagsanstalt, 1957 and 1964.

Trendelenburg, Friedrich Adolf. *Logische Untersuchungen* vol. I. Leipzig: Brockhaus, 1840; repr., Hildesheim: Olms, 1964.

Trendelenburg, Friedrich Adolf. *Unpublished Lectures on Psychology* (Ms. B 4) Trendelenburg-Nachlass. Staatsbibliothek zu Berlin, preußischer Kulturbesitz.

Trombetta, Antonio, see Antonio Trombetta.

Trutfetter, Jodocus, see Jodocus Trutfetter.

Tye, Michael. *Consciousness, Color, and Content*. Cambridge: The MIT Press, 2000.

Urban, Wolfgang. "Die '*via moderna*' an der Universität Erfurt am Vorabend der Reformation." In *Gregor von Rimini, Werk und Wirkung bis zur Reformation*. Edited by Heiko A. Oberman. Berlin: De Gruyter, 311–30, 1981.

Vasquez, Gabriel. *In Primam Secundae Divi Thomae*. Ingolstadt, 1612.

Verbeek, Theo. *Spinoza's Theologico-Political Treatise: Exploring "The Will of God"*. Aldershot: Ashgate, 2003.

Verbeek, Theo. "'Verum index sui et falsi': shinri-no 'Kihan' to 'Kijun'". *Spinozana: Supinoza Kyokai Nempo*, 4: 5–29, 2003.

Verbeek, Theo. "Wittich's Criticism of Spinoza." In *Receptions of Descartes: Cartesianism and Anti-Cartesianism in Early Modern Europe*. Edited by Tad Schmalz. London: Routledge/Taylor and Francis, 113–27, 2005.

Vernia, Nicoletto, see Nicoletto Vernia.

Vidal, Fernando. "Psychology in the Eighteenth Century." *History of the Human Sciences* 6: 89–119, 1993.

Volpi, Franco. "War Franz Brentano ein Aristoteliker? Zu Brentanos und Aristoteles' Konzeption der Psychologie als Wissenschaft." In *Die Brentano: Eine europäische Familie*. Edited by Kurt Feilchenfeldt and Luciano Zagari. Tübingen: Niemeyer, 134–38 and 143–45, 1992.

von Baer, Karl Ernst. *Über Entwicklungsgeschichte der Thiere*. Königsberg: Bornträger, 1828.

von Helmholtz, Hermann. *Handbuch der physiologischen Optik*. Leipzig: Voss, 1867.

Wallace, William A. "Traditional Natural Philosophy." In *The Cambridge History of Renaissance Philosophy*. Edited by C. Schmitt et al. Cambridge: Cambridge University Press, 209–12, 1988.

Watson, J.B. "Psychology as the Behaviorist Views It." *Psychological Review* 20: 158–77, 1913.

Watson, J.B. *Psychology from the Standpoint of a Behaviorist*. Philadelphia: Lippincott, 1919.

Wertheimer, Max. "Über Gestalttheorie." *Philosophische Zeitschrift für Forschung und Aussprache* 1: 39–60, 1925. An English translation "Gestalt Theory." Translated by Willis D. Ellis. In *Source Book of Gestalt Psychology*. New York: Harcourt, Brace and Co, 1938.

Williams, Bernard. *Descartes: The Project of Pure Enquiry*. London: Penguin, 1978.

Wilshire, Bruce. *The Primal Roots of American Philosophy: Pragmatism, Phenomenology, and Native American Thought*. University Park: The Pennsylvania State University Press, 2000.

Wittgenstein, Ludwig. *The Blue and Brown Books*. Oxford: Basil Blackwell, 1960.

Wittgenstein, Ludwig. *On Certainty*. Translated by Denis Paul and G.E.M. Anscombe. Oxford: Blackwell, 1969.

Wittgenstein, Ludwig. *Philosophical Investigations*. Translated by G.E.M. Anscombe. Oxford: Blackwell, 1953.

Wolfson, H.A. *The Philosophy of Spinoza*. 2 vols. Cambridge: Harvard UP, 1934.

Wollstonecraft, Mary. "Cave of Fancy." In *The Works of Mary Wollstonecraft*. Vol. 4. Edited by Janet Todd and Marilyn Butler. London: William Pickering, 1989.

Wollstonecraft, Mary. "Contributions to the *Analytical Review* 1788–1797." In *The Works of Mary Wollstonecraft*. Vol. 7. Edited by Janet Todd and Marilyn Butler. London: William Pickering, 1989.

Wollstonecraft, Mary. *A Vindication of the Rights of Woman*. In *The Works of Mary Wollstonecraft*. Vol. 5. Edited by Janet Todd and Marilyn Butler. London: William Pickering, 1989.

Wollstonecraft, Mary. *The Works of Mary Wollstonecraft*. Vol. 6. Edited by Janet Todd and Marilyn Butler. London: William Pickering, 1989.

Wood, Allen W. "Rational Theology, Moral Faith, and Religion." In *The Cambridge Companion to Kant*. Edited by Paul Guyer. Cambridge: Cambridge University Press, 394–416, 1992.

Wright, John P. and Potter, Paul. *Psyche and Soma: Physicians and Metaphysicians on the Mind–Body Problem from Antiquity to Enlightenment*. Oxford: Clarendon Press, 2000.

Wundt, Wilhelm. *Beiträge zur Theorie der Sinneswahrnehmung*. Leipzig: Winter, 1862.

Wundt, Wilhelm. *Outlines of Psychology*. Translated by C.H. Judd. Leipzig: Engelmann, 1902.

Zahavi, Dan. *Husserl und die transzendentale Intersubjektivität.* Phaenomenologica 135. Dordrecht: Kluwer Academic Publishers, 1996.

Zahavi, Dan. *Husserl's Phenomenology*. Stanford, CA: Stanford University Press, 2003.

Zeki, Semir. "Neural Concept Formation & Art: Dante, Michelangelo, Wagner." *Journal of Consciousness Studies* 9: 53–76, 2002.

Zupko, John Alexander. *John Buridan's Philosophy of Mind: An Edition and Translation of Book III of His 'Questions of Aristotle's De anima'*. Third Redaction, with Commentary and Critical and Interpretative Essay. Dissertation at Cornell University, 1989.

Author Index

Subject Index