Jon Miller
Editor

STUDIES IN THE HISTORY OF PHILOSOPHY OF MIND 9

Topics in Early Modern Philosophy of Mind

Springer

Topics in Early Modern Philosophy of Mind

STUDIES IN THE HISTORY OF PHILOSOPHY OF MIND

Volume 9

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.

For other titles published in this series, go to
www.springer.com/series/6539

Topics in Early Modern Philosophy of Mind

Edited by

Jon Miller

Editor
Jon Miller
Department of Philosophy
Queen's University
John Watson Hall
Kingston, ON K7L 3N6
Canada
jon.miller@queensu.ca

ISBN: 978-90-481-2380-3 e-ISBN: 978-90-481-2381-0

Library of Congress Control Number: 2009926323

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Contents

Acknowledgements

I wish to acknowledge the many individuals who helped bring this volume to fruition. First and foremost, there are the contributors themselves, all of whom were patient beyond expectation with an editorial process that seemed interminable. Next there are those who provided advice on how to structure the volume, including but not limited to Donald Ainslie, Deborah Brown, John Carriero, Alan Gabbey, Paul Hoffman, Steven Nadler, and Calvin Normore. Jordy Koffman was an excellent research assistant, going far beyond his copy-editing mandate. Finally, financial support of different kinds was given by Queen's University and SSHRC. Many thanks to all.

Contributors

SEAN ALLEN-HERMANSON is an assistant professor of philosophy at Florida International University.

TALIA MAE BETTCHER is an associate professor of philosophy at California State University, Los Angeles.

BRIAN JONATHAN GARRETT is an assistant professor of philosophy at McMaster University.

DON GARRETT is a professor of philosophy at New York University.

CHARLES JARRETT is an associate professor of philosophy at Rutgers University – Camden.

OLLI KOISTINEN is a professor of theoretical philosophy at the University of Turku.

JON MILLER is an associate professor of philosophy at Queen's University, Canada.

ANDREW PESSIN is a professor of philosophy at Connecticut College.

MARLEEN ROZEMOND is a professor of philosophy at the University of Toronto – Mississauga.

AMY M. SCHMITTER is an associate professor of philosophy at the University of Alberta.

ALISON SIMMONS is a professor of philosophy at Harvard University.

Introduction

Jon Miller

Commentators have long been struck by the connections linking contemporary accounts of the mind to those articulated by philosophers of the 17th and 18th centuries—the era commonly referred to as "early modern philosophy." For example, in his enormously influential *Language and Mind*, Noam Chomsky identifies two broad ways in which 20th century cognitive science unwittingly revisited themes much discussed during "the original cognitive revolution," as he describes the early modern period.[1] First, 20th century cognitive science became enamoured of "the view that had been crystallizing through the 18th century that properties 'termed mental' are the result of 'such an organical structure as that of the brain'."[2] Second, recent cognitive science concurred with the early modern thesis that "properties of the world termed mental may involve unbounded capacities of a limited finite organ …"[3]Writing as a cognitive scientist concerned to advance his vision of how the field ought to be reformed, Chomsky restricts the scope of his claims just to the congruencies between early modern concepts of mind and those of his discipline. If we drop that limitation, more connections between the early moderns and us are ready to hand.

For instance, Hilary Putnam credits early moderns with being the first to realize that "there was a serious problem about the relation of mind to material body."[4] While the relation was admittedly a problem for earlier thinkers, Putnam argues that because of the rise of modern physics in the 1600s, it intensified and assumed the form that we know today. The relevant development in physics, Putnam suggests, was the widespread recognition of the world as causally closed. Because "no body moves except as the result of the action of some *force*,"[5] then since force is a purely physical phenomenon, all bodily movement seems wholly due to the interaction between bodies. If that's the case, whence the mind? Not only did early moderns first formulate our version of the mind–body relation problem but also they proposed two solutions that have proven to possess enduring appeal. First, Putnam says, they held that mental events "could *parallel* physical events, e.g. events in the brain."[6] On this hypothesis, which Putnam ascribes to Spinoza among others, "mental events might actually be *identical* with brain events and other physical events."[7] Second, other early moderns insisted that whatever doubt might be cast on this possibility by the new physics, mental events

[1] Chomsky (2006, p. ix).
[2] *Ibid*. The quotations are from Priestley.
[3] *Ibid*.
[4] Putnam (1981, p. 75).
[5] *Ibid*.
[6] Putnam (1981, p. 76).
[7] Putnam (1981, p. 77).

"could interact with physical events. The mental events might actually be causing brain events, and vice versa."[8] The most sophisticated proponents of this position, such as Descartes, take the mind and the brain to be "an essential unity."[9] To be sure, Putnam thinks that recent philosophy of mind has hugely refined the problem as well as the solutions that he sees as first discovered by the early moderns. Nevertheless, Putnam argues, much of the agenda of contemporary philosophy of mind was set in the 1600s and 1700s.

Both Chomsky and Putnam describe, in distinct ways, the legacy of early modern accounts of the mind to our own theorizing about the mind. A different assessment has been provided by Richard Rorty. As Rorty reads early modern philosophy, Descartes invented a concept of mind that was radically different from what can be found in ancient and medieval philosophy. The properties of the Cartesian mind that Rorty deems significant are well-known: *inter alia*, the mind is a substance able to exist apart from the body and subject to its own principles.[10] More interesting is Rorty's view of the major challenge confronting the Cartesian mind, for Rorty thinks it is defined by its quest to grasp accurately the world that exists apart from it. While Descartes' successors modified his concept of mind in various non-trivial respects, Rorty argues that they still operated within the basic parameters that he established. Moreover, according to Rorty, the same was still at the time he wrote his book. Because philosophy from Descartes' time to the present thought of the mind in broadly Cartesian terms, philosophy throughout those centuries was dominated by the same epistemological project—viz., the project of determining how or whether the mind was successfully able to "mirror" nature. Now, Rorty's reading of history and his normative claims about the value of the philosophical traditions that he is criticizing are controversial and certainly wrong on many matters of detail. Regardless of whether every stage of his argument is correct, however, Rorty's take on early modern philosophy and its effects on contemporary philosophy is incontestably interesting. If it has any merit, for centuries early modern writings on the mind determined not only subsequent views of the mind but also the very enterprise of philosophy itself.

Given the similarities between their views and ours, not to mention the many debts that we owe to them, it is understandable that the writings of early moderns on the nature of the mind have been the subject of intense academic speculation in recent decades. The present volume marks a further contribution to this on-going endeavour. In the next part of this introduction, I will present a descriptive overview of the ten chapters. I will not make an argument as such. Rather, I want to help the reader by offering a gloss of the chapters, in hopes of clarifying their contribution to literature on early modern philosophy as well as contemporary philosophy of mind. After this overview is complete, I shall conclude by pulling out some prominent themes.

[8] Putnam (1981, p. 76).
[9] Putnam (1981, p. 77).
[10] See the list on p. 35 of Rorty (1979).

Andrew Pessin begins the volume by grappling with a question faced by generations of Descartes scholars: What is immediately before the mind when it senses an object external to itself? To borrow an example from Pessin, what does the mind itself directly sense when it sees the sun? According to one school of thought, Descartes held that the mind senses the object's secondary qualities, such as colour, and these secondary qualities do not resemble its primary qualities, such as extension. Because the mind does not sense the object's primary qualities, it can only infer them on the basis of its experience of the secondary ones. Pessin opposes this interpretation. Defining his terms with care and precisely locating his reading in the relevant primary and secondary literature, Pessin's argument unfolds in two broad stages. The bulk of his paper is devoted to making the case that the object of sensation can be literally in the senses. If this is true, then there is no re-presenting of the object in sensation, since sensation is directly of the object. Starting in Section 8 of his paper, Pessin complements his analysis of the mechanics of sensation by examining the ontological status of the phenomena of experience. He contends that the phenomena of experience—the secondary qualities that we are aware of when we have experiences—have no being. Even though we are aware of colours and other sensory qualities, Pessin argues that they don't stand as objects of sensation, for Descartes thought they were reducible to the motions of the surfaces of the objects being sensed. Since secondary qualities such as colour have no being, they cannot be the objects of sensation, for something which is nothing cannot be sensed.

Pessin's is the only paper in the volume exclusively focused on Descartes. Brian Jonathan Garrett's is the first of several that situate Descartes in a broader context. Interested in the status of wonder in philosophy, Garrett asks why Descartes and Malebranche are suspicious of the contributions that wonder can make to philosophical inquiry. The question is fair; after all, the dictum that philosophy begins in wonder goes back to the Greeks. Moreover, the early modern period was a time fraught with discoveries and inventions that prompted wonder among commoners and cognoscenti alike. As Garrett reads Descartes and Malebranche, they bifurcate the role of wonder in philosophy. In the moral domain, wonder is legitimate, as it alone is able to reveal to us certain essential truths, especially the truth that our wills are free. On the other hand, Descartes and Malebranche would entirely purge metaphysics and science of wonder. The basic problem is that wonder doesn't teach us anything that we can't learn by other means and it stands in the way of the serious and critical thinking that is constitutive of metaphysical and scientific inquiry. To make and sustain these points, Garrett takes us on an entertaining journey through an array of 17th century literature on wonder.

Where Garrett is moved to study Descartes and Malebranche's views on wonder because of a wider interest in the history of wonder, the next paper in the volume approaches its subject from the opposite direction. Contemporary theorists of the mind such as Chomsky have argued that our understanding of the mind can be enhanced through the study of early modern analyses of mind. Following this

suggestion, Sean Allen-Hermanson looks at the views of Desgabets, a very peculiar early modern. Desgabets is peculiar and interesting because he attempts to blend Cartesian rationalism with empiricism. Or so the majority of scholars have thought until recently. As Allen-Hermanson tells us, however, Monte Cook has mounted a forceful challenge to this reading, arguing instead that Desgabets rejected core tenets of empiricism. Allen-Hermanson uses Cook's challenge to mount a re-assessment of Desgabets' empiricism. To conduct this re-assessment, Allen-Hermanson compares the commitments found in Desgabets' writings against eight empiricist credos. These credos include the thesis that ideas resemble sensations, or that all knowledge is *a posteriori*, derived from experience. The search undertaken by Allen-Hermanson reveals that Desgabets embraced versions of all eight propositions. At the same time, Allen-Hermanson does not conclude that Desgabets was a straightforward empiricist, since he continues to embrace core Cartesian doctrines, such as Descartes' conception of sensible qualities. Rather, he thinks of Desgabets as conceiving of higher mental functions in rationalist terms, de-emphasizing sensation in favour of innatism, while taking lower mental functions to work as the empiricists described. So, in the end, Allen-Hermanson defends the traditional view of Desgabets as a unique figure in the 17th century straddling both sides of the rationalist–empiricist divide.

If Brian Garrett and Allen-Hermanson grapple with an issue or a figure that may seem *recherché*, Don Garrett poses a question that is well known: is nothing simultaneously thinking and extended? Early modern opinion was divided. On one side, philosophers such as Malebranche, Leibniz and Kant followed Descartes in holding the affirmative. On the other, Spinoza, Locke and others maintained the contrary, that one-and-the-same thing could be both thinking and extended. In his paper, Garrett gets at the heart of the debate by asking what Spinoza and Locke would have made of Descartes' claim that no thinking thing is also extended. According to Garrett, there are two chief arguments given by Descartes for his claim, one based on the separability of mind and body, and the other on the divisibility of body versus the indivisibility of mind. In characteristic fashion, Garrett meticulously reconstructs each argument, identifying the major assumptions Descartes makes and then explaining the process of reasoning which leads up to the conclusion. After the argument is in place, Garrett proceeds to determine exactly which assumptions or inferences would be resisted by Spinoza and Locke. There is much that Garrett's readers will find of interest in his paper. Some will gravitate toward his reconstruction of the arguments, which they may or may not find convincing. Others will be drawn to the conjectural responses offered on behalf of Spinoza and Locke. Garrett ends his paper by connecting the early modern debate that he describes to a problem central to contemporary philosophy of mind: namely, the problem of how something material can be conscious. Like Allen-Hermanson, Garrett thinks new light may be shed on this old problem by revisiting the arguments of the early moderns, especially Spinoza.

In her paper, Alison Simmons engages with a question that is just as enduring and just as grand: what is truly essential to the mind? When answering

this question, one can tap into a plethora of mental powers, faculties or attributes. Interestingly, there are discernible trends regarding the answers given. It is commonplace nowadays to emphasize beliefs, desires, reasons for action—all the things that philosophers sum up with the word "intentionality"—and consciousness as hallmarks of the mind, but historically, this has not always been true. Specifically, several recent commentators have argued that Malebranche's conception of the mind stands out as one of the few which not only de-emphasized intentionality but also denied it any place in the mind's essence. Simmons doesn't deny that Malebranche's conception of the mind is remarkable but she does wish to convince her readers that what makes it remarkable is not that it eliminates intentionality. Granted, what Malebranche makes of intentionality is unusual, for he scrubs it of all representationality. To say that, however, is not necessarily to say that thoughts are *of* nothing. As Simmons reads Malebranche, he holds that when the mind adopts the intentional stance, it is thinking of something. However, what it is thinking of is not intrinsic to the thought itself but rather the content of the thought is fully constituted by actually existing things outside the mind. As Simmons puts it, "Malebranche conceives the mind's intentionality as an extrinsic relation between mind and extramental objects."[11] It is not, however, Malebranche's theory of intentionality that truly distinguishes him from his peers. According to Simmons, what's most distinctive about Malebranche's conception of the nature of mind is the importance assigned to consciousness. Unlike his peers, Malebranche denigrates consciousness as a form of cognition that is inferior to other forms of cognition. Simmons ends her chapter by suggesting that Malebranche may be the first in the history of philosophy to notice and expound upon the epistemological and methodological difficulties surrounding consciousness.[12]

Depending on a philosopher's views on the nature of mind, questions may arise as to what unifies the mind. In his paper, Charles Jarrett explores the difficulties surrounding Spinoza's account of the mind and its unity. Some of the difficulties are due to Spinoza's conception of the mind as the idea that God has of a specific human body. On this view, my mind is unified insofar as my body is unified. Of course, this account of the mind's unity only serves to push back the question, for we should now ask what accounts for the unity of my body. As Jarrett tells us, Spinoza has an answer to this question, given (among other places) in the digression into physics found after Proposition 13 of Part II of the *Ethics*. The unity of the body consists entirely in a ratio of motion-and-rest that exists among that body's parts. So long as the ratio remains, the body remains the same. If the body's unity is bound up with the preservation of a specific ratio of motion-and-rest among its parts, and if the mind's unity is parasitic upon the body's, then it might seem to follow that Spinoza thought of the mind's unity in similarly mechanistic terms as he did the body's. While Jarrett ultimately defends the

[11] Simmons p. xxx.

[12] Though incomplete (for example, it entirely omits Malebranche's possible contributions), Heinämaa et al. (2007) provides an excellent overview of the history of consciousness.

inference, he finds it necessary first to take note of some complicating factors. One complication is that the mind is described as the "idea *of*" the body. This description makes the mind inherently representational—it represents the body—and since representational entities are not obviously identical in the relevant respects with mechanistic ones, Jarrett thinks this consideration alone ought to give us pause before understanding the mind's unity in mechanistic terms. A different complication stems from Spinoza's views on final causes and teleology. A functionalist account of the mind's unity would hold that the mind's unity is ascribable to that mind's possessing a single primary function or goal. Jarrett explores the case for taking Spinoza to subscribe to such an account, taking us through the murky texts which relate to final causation and teleology. In the final section of his paper, Jarrett bravely travels into the most controversial texts in all of Spinoza's corpus, wherein he presents his views on immortality. Jarrett argues that these texts support a twofold account of the mind's unity: one for when the mind exists in a certain time and place, and another for when it exists atemporally and aspatially. He concludes by connecting this Spinozistic thesis of mental unity to Kant's distinguishing between the unity of a thing *qua* phenomenal entity versus noumenal entity.

The next chapter in the volume picks up some of the same issues. Olli Koistinen's starting point is the widely-held view that Spinoza subscribed to a "bundle theory" of the mind. As Michael Della Rocca explains, "On the bundle theory, the human mind is not a thing that is distinct from its ideas. Rather, the mind is just a collection of ideas that bear a certain relation to each other."[13] Koistinen concedes that there are firm grounds for taking Spinoza to have held that the mind is a bundle of ideas. Given that the "order and connection of ideas is the same as the order and connection of things" (*Ethics* IIP7), then since the body is a collection of so many simpler bodies standing in a certain specific relation to one another, the mind must likewise be a collection of simpler ideas standing in a certain relation to each other. At the same time, Koistinen argues, even if bundling is a necessary part of any mind, it is far from sufficient, for even if the mind as an abstraction is a bundle of ideas, it doesn't follow that the mind is *some particular person's* mind. We can conceive of a coherent bundle of ideas as forming a coherent whole without forming the mind of any recognizable person. So, Koistinen asks, besides being a coherent whole, what else is required of a bundle of ideas for it to be regarded as a mind? Koistinen argues that the missing element is the principle of action for the whole: a mere bundle of ideas lacks a principle of action; a genuine mind has one. As Koistinen notes, there is a hint of paradox to the suggestion that some minds are sources of action, for every mind is *ipso facto* finite and overwhelmed by the power of the one substance. To develop and, ultimately, defuse this paradox, Koistinen invokes Kant's conception of the agent as *actus originarius*, the being whose action is not determined by anything outside itself. As Koistinen reads Spinoza, the power of action which differentiates

[13] Della Rocca (1996, p. 41).

random bundles of ideas from genuine minds or, as Koistinen prefers to call them, true "selves" is unitary or non-composite. So while Spinoza may be said to have a bundle theory of the mind, he did *not* have a bundle theory of the self. Koistinen concludes by applying his analysis of Spinoza's conception of the mind versus the self to the problems of immortality, also tackled by Jarrett. Koistinen argues that at least some of the problems are solved by the distinction between mind versus self.

In the next chapter, Marleen Rozemond returns to the central question of Don Garrett's paper: can matter think? The difference in philosophers that Rozemond addresses—Clarke and Collins versus Garrett's Descartes, Spinoza and Locke—commits her to a difference in the orientation of her problematic, for while Descartes, Spinoza and Locke take positions on whether there can be thinking and material *substances*, Clarke and Collins argue over whether consciousness can *emerge* from material qualities. As Rozemond says, insofar as current philosophy of mind has abandoned talk of substances in favour of talk of mental and physical states, the Clarke/Collins exchange will seem more topical than the Cartesian dialectic. As Rozemond retells the Clarke/Collins exchange, it begins with Clarke's version of the so-called "Achilles Argument." Though there are many different versions of the argument,[14] all of them infer the simplicity of the mind from the unity of thought or consciousness. After going through Clarke's version of the argument, Rozemond presents Collins' attempts to clarify the exact import of the argument, especially on the question of the precise conditions for consciousness. While these clarificatory probes are interesting, Rozemond thinks that Collins' true advance comes in the form of a radically different conception of the relationship between mind and matter. Collins grants that mind and matter have essentially different natures; yet, he denies that this makes it impossible for mind and matter to stand in a causal relationship to one another. Just as a musical instrument can produce harmony or a clock can tell time, so the mind can emerge from matter of the brain. Individually, the parts of the brain are unable to produce consciousness and thought, but collectively, working in concert, they do. As Rozemond makes plain, Clarke stoutly rejects Collins' emergentist conception of the mind. One problem is the lack of motivation. All of the examples that Collins provides of distinctly new qualities emerging from older ones different in kind from them are not truly inherent qualities. The musical harmony produced by an instrument is not really inherent in it, for the harmony essentially involves a listener. Consciousness, however, is an inherent quality: it belongs to the very states allegedly producing it. Because Collins has not given any legitimate example of an inherent quality that is also emergent, Clarke contends that his entire enterprise is unmotivated. Toward the end of her paper, Rozemond provides a different perspective on the dispute between Clarke and Collins. It concerns the limits of our knowledge. According to Rozemond, Clarke is an optimist about the future prospects for inquiry into the mind. He thinks that humans can or perhaps

[14] For more on the argument, see the collection of papers in Lennon and Stainton (2007), which includes another essay by Rozemond on the Clarke–Collins correspondence.

already do know enough about consciousness to know that it cannot be identified with motion or any other material quality. Collins is more diffident. It is entirely conceivable, says Collins, for there to be qualities of matter about which we are and will remain ignorant. For this reason, we simply cannot completely discount the possibility that the mind can emerge from those qualities.

Rozemond deals with texts that are not widely read. By contrast, the epigraph and focal point of Talia Mae Bettcher's paper is one of most famous passages from Hume's *Treatise*. "[W]hen I enter most intimately into what I call *myself*," Hume wrote, "I always stumble upon some particular perception … I never catch *myself* at any time without a perception …"[15]These words express Hume's views on the nature of the self, views which have been enormously influential. While many commentators have analyzed Hume's concept of the self on its own terms or in relation to other early moderns, especially Locke, Bettcher takes the novel approach of seeing it as a response to Berkeley. *Prima facie*, this may seem unpromising, for while Berkeley distinguishes between the self and ideas, Hume does not in the body of the *Treatise*.[16] This distinction is of vital importance, for it entitles Berkeley to something that Hume denies himself: namely, the notion of a self, standing above-and-beyond all its perceptions or ideas. Bettcher provides several reasons for taking the first appearances to be misleading. *Inter alia*, she notes the temporal indexicality of Hume's remarks. They concern what he observes of himself *in the present*. Even if he should fail to "catch" himself "at any time," it certainly doesn't follow that he would also fail to catch himself after that time. For this and other reasons, Bettcher urges us to suspend any initial reservations we might have about connecting Hume to Berkeley. Bettcher does not argue that Hume is a closet Berkelian, for even the re-examination of the relevant texts that she undertakes—a re-examination avowedly undertaken to reassess the extent of the Berkelian hypothesis appropriated by Hume—yields key differences between the Scot and the Irishman. At the same time, Bettcher argues in her final section, Hume and Berkeley agree on one key point, a point on which Hume has more in common with Berkeley than Locke. Hume held that introspection only reveals awareness of perceptions. Since existence is not a subject for perception, one cannot perceive one's existence. Because Locke held that one perceives one's existence over-and-above one's perceptions, the Lockean conception of the self is flatly incompatible with the Humean one. By contrast, while Berkeley agreed with Locke that one is conscious of one's existence, he did not think that one *perceived* one's existence. Rather, he took the self to be an unperceived perceiver. Since he didn't establish the self and its existence as the object of perception, there is no incompatibility between the Berkelian hypothesis and Hume's.

[15] *A Treatise of Human Nature*, I.iv.6.3.
[16] As Bettcher notes, however, Hume does acknowledge the distinction in the Appendix to the *Treatise* (see Bettcher, p. xxx).

Hume's views on the self also feature in the final essay of this volume. As Amy Schmitter notes, Hume endorses several propositions that seem at odds with each other.[17] On the one hand, Hume holds that (1) passions such as pride and humility have an object, (2) the object of all passions—even those that are contrary to one another such as pride and humility—is the same, and (3) the object of the passions is the self. On the other hand, Hume seems to argue for (4) a deflationary account of intentionality and (5) a conception of the self which makes it nothing more than a bundle of perceptions. It is not obvious how (1)–(3) can be reconciled with (4)–(5), though Schmitter thinks reconciliation is possible. The key, she argues, is Hume's account of intentionality *per se*. Now, Schmitter acknowledges that much of what constitutes intentionality for Hume will be similar to the intentional as conceived by his empiricist predecessors. The notion of perceptual contact between phenomena and the mind, so important to Locke, is accepted by Hume as key to accounting for what it is to be a mental object. Moreover, Hume finds value in Berkeley's copy-relation thesis, whereby intentionality is partially determined by the ability of the mind to copy conceptual content into its ideas. This copy-relation thesis helps to explain how a perception comes to have content. All the important inheritances from Locke and Berkeley to Hume notwithstanding, Schmitter argues that Humean intentionality is deeply original. Unlike Locke and Berkeley, Hume subscribed to a holistic conception of the intentional. Schmitter quotes Hume as writing, "'tis a maxim, which I have just now establish'd, and which is absolutely necessary to the explication of the phaenomena [of the passions], 'That 'tis not the present sensation or momentary pain or pleasure, which determines the character of any passion, but the general bent or tendency of it from the beginning to the end."[18] Once we understand the holism of Humean intentionality, Schmitter argues, we can revisit our understanding of (4) and (5). While the non-holistic elements admitted by Hume into intentionality are indeed minimal, the holistic ones have the effect of rendering intentionality more robust than one might otherwise think. Also, while the self is a bundle of perceptions, the contents of one's perceptions include elements that track through time, so that there is more continuity to the self than often supposed. Once (4) and (5) are understood correctly—Schmitter does not jettison them altogether—the seeming incompatibility of them to (1)–(3) dissolves. All passions can have the self as their object because passions include their histories in themselves, as does the self whose passions they are.

The title of this volume is reflective of its contents, for there is no single topic that the contributors were charged with examining. Nevertheless, certain features or

[17] To understand those propositions, a pair of definitions will help. When Hume speaks of "passions," he can be taken to mean "emotions." Also, when Hume speaks of the "object" of the passions, Schmitter says, he is referring to what the passion is about or what it is directed towards. To use the word we encountered earlier in Simmons' essay, the "object" of a passion is its intentionality.

[18] *A Treatise of Human Nature*, II.ii.9, 384-385.

themes emerge through the course of the ten chapters that make the whole more than just the sum of its parts. To conclude this introduction, I will draw attention to three of them.

First, contributors were asked to address a range of philosophers, from those who continue to engage us to those whom we have forgotten. No explanation need be given for including the usual suspects (such as Descartes, Spinoza and Locke) but a few words ought to be said for expanding the parameters to more obscure figures (Desgabets, Clarke and Collins). The rationale here is twofold. First, it is essential to the project of recovering the dialectic of early modern philosophy of mind, since both the famous and the now-forgotten made contributions to the discussion. In addition, a great deal of interesting philosophy can be found in the works of men like Desgabets. So, this volume contains papers on a number of different philosophers, in the expectation that it will teach us about philosophy of mind in the 17th and 18th centuries as well as philosophy of mind as such.

Second, there is the reappearance of certain concepts and problems in different chapters. The notion of the self, the nature of intentionality, the mystifying question of whether extended beings can think, representationalism—each of these is dealt with by at least two papers. The reappearance of the same issue in different papers is an indication of the importance of that issue, both in the eyes of contemporary scholars as well as in the thought of the early moderns. It is also an educational opportunity, for it is illuminating to compare different philosophers' views on the same issue, especially when the philosophers belong to the same historical era.

Of course, it is equally illuminating to test philosophers' views against our own best estimation of the truth. And this leads me to the final feature of the volume. Many of the contributors explicitly assessed the merits or demerits of the historical figures or ideas they were discussing. For example, Don Garrett ends his paper with a spirited defence of Spinozistic panpsychism, on the grounds that it holds greater promise than the alternatives to explain how one-and-the-same thing can be both thinking and extended. Or, to cite another instance, Simmons thinks that while Malebranche's theologically-driven account of the mind will be largely unacceptable to our secular age, those very same religious presuppositions should make it so unusual and hence interesting to us. Not all contributors are so positive about the prospects of the theories before them. Thus, Jarrett, after establishing that Spinoza takes the mind to consist of two parts, one eternal and the other not, glumly concludes that it is "not easy to see" how they can be unified into a single entity.[19] It is a sign of the respect these philosophers have for those whom they are discussing that they would not only recount their ideas but also submit them to rigorous scrutiny. They surely hope their own readers will do the same to them.

[19] Jarrett (p. 213).

Mental Transparency, Direct Sensation, and the Unity of the Cartesian Mind

Andrew Pessin

Introduction

An old question continues to rankle: Does Descartes hold that external objects are cognized only indirectly, by means of our directly cognizing some form of representative stand-in? Does he subscribe, in other words, to a "veil of perception" account of cognition?[1] The view that he does, in one form or another, has been defended by many scholars,[2] as has the view that he does not,[3] in one form or another, with the latter generally addressing "primary" qualities. In this paper I examine the following (style of) argument, which has provided much impetus for the "veil" camp, particularly with respect to "secondary" qualities:

> (A) Cartesian sensory qualities—in particular, "secondary" qualities—do not resemble the real (or "primary") physical qualities generally causing sensory perception, and so must be construed as mental entities. During sensation we are aware of the former and not the latter. Therefore, either we are not aware of the physical qualities at all, strictly speaking, or, we are aware of them only indirectly, by being aware of the sensory qualities which are their representative stand-ins. In either case sensory qualities, or ideas thereof, constitute a "veil" screening the mind from direct sensory contact with the external world.[4]

Though (A) refers to Cartesian secondary qualities, in fact it instantiates a very general, and centuries-old, style of argument, some instances of which apply to sensation, some to intellection, and some to both. First, one notes that "what" we cognize, or "what" we are aware of during sensation or intellection, differs from what is really "out there;" then one infers that what we cognize is a distinct entity

[1] The "veil of perception" phrase comes from Bennett (1971, p. 68).

[2] Kenny (1968), Hacking (1975), Rorty (1979), Maull (1980), Arbini (1983), Chappell (1986), MacKenzie (1989, 1990), Wilson (1978, 1994), Hoffman (2002), Clemenson (2005). Some of these authors have "mixed" views: for example, MacKenzie (and perhaps Arbini) rejects the "veil" for cognition of "primary" qualities but accepts it for "secondary."

[3] Lennon (1974), O'Neil (1974), Aquila (1974–5), Costa (1983), Yolton (1984), Cook (1987), Nadler (1989), and Alanen (2003).

[4] For one version of (A), see Wilson (1994, p. 19). Cf. also Costa (1983, p. 546); McRae (1975); Yolton (1975, pp. 384–385).

J. Miller (ed.), *Topics in Early Modern Philosophy of Mind*, Studies in the History of Philosophy of Mind 9, 1–37.

(or property) from what is out there; then one concludes that what we cognize veils us from what is out there, so that cognition of the physical world is indirect at best.[5] There is no question that Cartesian texts often support versions of these arguments or their conclusion, and so the veil reading is textually well-founded.[6]

The problem, though, is that all of the key terms are theoretical in nature, and so open to multiple interpretations. That is certainly true of such terms as "direct" and "indirect" with respect to cognition, as well as of "ideas," "representation," and "perceptions." But as is clear from (A), much also turns on how we understand the "object" of cognition, of "what" one is cognizing or aware of. Descartes' tendency to understand mentality or thought in terms of conscious awareness, and to recognize that such awareness always has an object, i.e., is "of" something, makes this a pressing question.[7] Unfortunately, his simultaneously both reflecting and attempting to reject scholastic views about cognition (as we'll see) also makes it a difficult question. Immediately there are competing candidates for being "what" is cognized: external objects, forms or essences, mental entities *representing* objects, the mental act of cognizing itself, or aspects thereof. Complicating the discussion is the fact that one must distinguish the possible *causal* intermediaries linking an external object to an act of cognition from possible *cognitive* intermediaries, a distinction Descartes is not always careful to make explicitly. Similarly, recent scholars have been influenced by Wilson's (1990) distinction between "presentational" and "referential" mental content in Descartes (as we'll see), which itself introduces a deep ambiguity into our question. And finally, the possibility that for Descartes something might be the "object" of a sensory state—what we are aware of—without our knowing *that* it is so makes the question very close to impossible to answer.

Yet I will try to answer it. I will argue that Cartesian sensory cognition is direct, and it is so because its objects, "what" we sense or are aware of during sensation, are (despite all appearances, literally) mechanical properties of the physical world. Towards the end I will also suggest that, in fact, Descartes was an eliminativist about what we might call qualia: that Cartesian phenomenal colour (for example) is, ultimately, *nothing at all,* and is, therefore, not "what" is seen during sensation.

---

[5] Other examples of this style make reference to: (i) Non-existing objects. (For medieval discussion of this form of argument, see Pasnau (1997); it was also debated by Malebranche, Arnauld, Reid, etc.; and it has recently been invoked by Chappell (1986, p. 185) and Wilson (1994, pp. 17ff.)) (ii) Illusions, errors, relativity of perception. (See Pasnau (1997) for medieval references; cf. Locke and Berkeley on relativity of perception arguments, etc.) (iii) "Presence to mind." (See the same sources. Arnauld was particularly critical of Malebranche's use of such an argument (*True/False* Ch. 4, *VFI* 38.190–197).)

[6] Just one example: *6th Med.*, AT VII.75, CSM II.52.

[7] *2nd Replies*, AT VII.160, CSM II.113; *Principles* I.9, AT VIIIA.7, CSM II.195; *4th Replies*, AT VII.246, CSM II.171.

1 Preliminaries

Some terminology. I follow Simmons (1999) in understanding Cartesian sensations to "include all those obscure and confused modes of mind that arise from the union … of mind and body: conscious experiences of pain, tickling, hunger, thirst, light, colours …" etc., and to "constitute the qualitative character … of human experience"[8] (*pace* the eliminativism to be discussed later); I follow Simmons (2003) in noting that "Descartes uses the terms 'sensory idea,' 'sensory perception,' and 'sensation' more or less interchangeably."[9] A sensation of (say) yellow, or a sensory idea of yellow, is a mode of the mind, or, more precisely, is the state (or act) of sensing yellow. I shall generally restrict my use of "sensation" to the so-called secondary qualities. I shall use "conception" or "intellection" to refer to the states of the intellect, themselves also modes of mind; the intellectual idea of (say) the sun shall correspondingly be the state of conceiving the sun.[10] I'll use "cognition" to remain neutral between sensation and intellection. I will distinguish between the "representational content" or "object" of a cognitive state and its "qualitative character," where the former refers to what the state represents or is "of," i.e., "what" is cognized, and the latter refers, roughly, to "what it's like" for the cognizer to be in that state. The precise relationship between them I leave open for now. An object or property x is cognized "directly" where x is the object of cognition; x is cognized "indirectly" where a cognitive state in some way is referred to x but where something distinct from x is the object of the state.[11] Using as our primary example an instance of Fred seeing the sun, let Fred's state of sensing the yellow light be "sensation-y," and the primary qualities of the sun distally initiating the causal sequence eventuating in y be "motion-m." *Prima facie*—but only *prima facie*—we might say that y is "of" yellow; the object of Fred's sensation, "what" he sees, is yellow (or yellowness). To determine whether Cartesian sensation of the sun is direct, then, we must determine whether the object of y, what Fred sees, is also (or in fact) m. If so—as I shall argue—then Cartesian sensation is direct.

---

[8] Simmons (1999, p. 347).

[9] Simmons (2003, p. 551). See *Principles* I.66, AT VIIIA.32, CSM I.216, and IV.189, AT VIIIA.316, CSM I.280.

[10] That is, the sun insofar as we are thinking about it and not currently sensing it, and in particular thinking about it on the basis of astronomical reasoning (*3rd Med.*, AT VII.39, CSM II.27). Such an idea would involve clear and distinct ideas of extension and its modes.

[11] There are many other senses of "direct" and "indirect" to be found in discussions of cognition: direct cognitions might be non-inferential, certain or infallible, first-person privileged, etc. We'll also see a couple of other senses in the course of this paper. But my primary interest is in the sense just characterized. (For various senses of "(in)direct," see Nadler (1989), Tipton (1992), and Hoffman (2002).)

2 Mental Transparency, Objective Being, and Direct Intellection

Unfortunately, my case is immediately made more difficult by Descartes' doctrine that the mind is "transparent" to itself, that is, aware or conscious of all its own occurrent states.[12] For our purposes, transparency entails at least that when sensing yellow, Fred will be aware not only of the yellow but also of his own state of sensing. This suggests that his own state of sensing might itself be an object of cognition, perhaps even its *own* object. But this, in turn, could go far to underwrite argument (A): "What" we're aware of, in sensation, is (only) our own mental state. In sensing yellow perhaps Fred is merely aware of (say) the qualitative character of his own state of mind.

In fact, things are even better for (A). Descartes himself distinguishes between a "direct" or immediate awareness of one's mental states[13] and one due to reflection, viz. by means of another mental state.[14] The direct awareness appears to be a primitive feature of *every* mental state; it may also be the source of whatever first-person epistemic privilege the agent enjoys with respect to his own mental states. The reflective awareness is due to the intellect, which can (but need not)[15] consider, in various ways, its mind's own modes.[16] So, by transparency, we are "directly" aware of our own mental states, just as (A) asserts;[17] our awareness of the external world might thus seem to be indirect at best.

---

[12] *4th Replies*, AT VII.246, CSM II.171; To Mersenne, 1/28/41, CSM III.172. More generally, Descartes sees the essence of the mind to be thought, and holds that conscious awareness is essential to thought: *2nd Replies*, AT VII.160, CSM II.113; *Principles* I.9, AT VIIIA.7, CSM I.195. Transparency so construed is a fairly weak notion, with no requirement (say) for infallibility, or incorrigibility, or even the "completeness" of our self-knowledge, etc. For discussion, see Rozemond (2006).

[13] *6th Med.*, AT VII.75, CSM II.52; To Mersenne, 1/28/41, AT III.295, CSM III.172. Cf. *Passions* I.19, AT XI.343, CSM I.336; *6th Replies*, AT VII.422, CSM II.285; *7th Replies*, AT VII.559, CSM II.382; *3rd Replies*, AT VII.181, CSM II.127; *2nd Replies*, AT VII.160, CSM II.113.

[14] *Discourse* III, AT VI.23, CSM I.122; *Burman*, AT V.149, CSM III.335; *5th Replies*, AT VII.358–359, CSM II.248; *3rd Replies*, AT VII.182–183, CSM II.128; For [Arnauld], 7/29/48, AT V.220–221, CSM III.357. Cf. Arnauld, *True/False*, Ch. 6, 71; *VFI* 38.204. For discussion of Descartes' distinction between immediate awareness and that due to reflection, see McRae (1972), Wilson (1978, p. 150ff.), Vinci (1998, p. 39ff.).

[15] *Burman*, AT V.149, CSM III.335. For [Arnauld], 7/29/48, AT V.220-221, CSM III.357.

[16] Regress is avoided by holding that one is primitively aware of the intellectual act underwriting reflective awareness of other acts. Hobbes raises the regress worry at *3rd Objections*, AT VII.173, CSM II.122; for discussion, see Radner (1988, p. 442ff.), Vinci (1998, p. 23ff.) (cf. also note 92).

[17] Indeed, some Cartesian texts suggest that we are directly or immediately aware only of our own mental states: *3rd Replies*, AT VII.181, CSM II.127; *6th Med*., AT VII.75, CSM II.52; *2nd Replies*, AT VII.160, CSM II.113. Note that the examples Descartes gives in these texts of what is "within us" or "immediately perceived" are mental states or operations. Consider, too, *Principles* I.8, AT VIIIA.7, CSM I.195.

Consider (A) applied to intellection. Suppose Fred is contemplating the astronomical idea of the sun. That is, Fred is thinking of the sun; the sun is the object of his thought. His act of thinking of the sun is a mental state, while the sun is a physical object; the two cannot be identified. Given transparency, Fred is directly (non-reflectively) aware of his thinking of the sun. But if he is directly aware of his thinking of the sun, and that act of thinking is not identical to the sun itself, then he would seem at best only indirectly aware of the sun itself. His cognition of an external object is in this way mediated by his awareness of his own mental state. And if this is so in an 'intellectual' case, then all the more so in sensation, where (say) the yellowness Fred senses does not even resemble what's in the sun itself. What Fred is aware of, in sensing the sun, simply cannot be identified with the sun. We have our veil.

Yet despite the appeal of such reasoning—and apparent support in Cartesian texts—I believe it should be resisted. The sense here in which Fred's awareness of his own mental state is "direct," I will argue, need not conflict with the sense in which Fred's awareness of the *sun* is direct.

(1) First, note that despite the epistemic priority enjoyed by our awareness of our states over our awareness of external objects,[18] Descartes never seems tempted by the obvious consequence of a "veil" view, viz. that we don't actually cognize external objects at all, but only our own mental states.[19] This suggests, I think, that despite his non-reflective awareness of it, Fred's state is *not*, after all, an object of cognition, properly speaking, or at least not an object ultimately to be distinguished from whatever *is* the object of cognition;[20] and if not, it is not in competition with the sun as being the object of Fred's cognition.

(2) When Fred thinks of the sun, there occurs in his mind at least

(a) An awareness of the sun, and, given transparency
(b) Direct awareness of his awareness of the sun

In calling (b) "direct" we are stressing that it does not require any additional mental act beyond (a). But that means that, in fact, Descartes would not distinguish (a) and (b) but rather *identify* them, as many of the texts cited in note

[18] Such as the infallibility about the former lacking about the latter. Or as Wilson (1994) notes: "… [O]ur knowledge that we actually perceive any bodies is dependent upon, and derived from, the epistemologically prior affirmation that we have in our minds ideas 'of them'," (p. 17) and "… Descartes can know all about the objective as well as the formal aspects of his ideas, without yet knowing whether any extra-mental entities exist." (p. 18) (Cf. *Principles* I.8, AT VIIIA.7, CSM I.195.)

[19] Neither with respect to intellection, nor to sensation. He is concerned in the latter first with whether they are caused by external bodies hence reliable indicators of their existence, not with whether they are "of" external bodies (cf. *6th Med*.); and he is further concerned (as we'll see) not that we do not sense physical things, but that we do not sense them *intelligibly* or *accurately*. For example: *Principles* I.70, AT VIIIA.34, CSM I.218; *Passions* I.23, AT XI.346, CSM I. 337.

[20] Non-reflective awareness, then, is neither sensory nor intellectual, hence not cognitive (in the sense in which I'm using the term). Cf. Costa (1983): "Nor do I think that Descartes considered conscious awareness to be anything akin to seeing." (p. 547)

13 suggest. But if (b) is not distinct from (a), then there is only one mental state: Fred's awareness of his awareness of the sun just *is* his awareness of the sun. If so, then (b) would not suggest that Fred's awareness of the sun is indirect in our sense.

(3) Consider, too, Descartes' tendency to explicate transparency in terms of the awareness of the *object* of a mental state. He writes, for example: "*Idea*. I understand this term to mean the form of any given thought, immediate perception of which makes me aware [*conscius*] of the thought."[21] "Form" here refers to those features of a mental state which make it both the kind of state it is (such as a conceiving, a fearing, etc.),[22] as well as provide its object.[23] Here, the awareness of his mental state described by transparency is attributed to the immediate perception of that which provides the state's object. This suggests that Descartes does not distinguish awareness of one's mental state from awareness of the object of that state. But then, again, Fred's awareness of his awareness of the sun in (b) above might just *be* his awareness of the sun. If so, then (b) would not entail that Fred's awareness of the sun is indirect.

(4) Consider, too, one of the well-known ambiguities in Descartes' use of the word idea: "'Idea' can be taken materially, as an operation of the intellect …. Alternatively, it can be taken objectively, as the thing represented by that operation."[24] When Descartes speaks of "immediate perception" of "ideas," then, that might (depending on context) refer to the direct awareness of our mental states, as (A) suggests, but it also might refer to the direct awareness of the objects of those states. Transparency, again, seems to apply equally well to our mental states and their objects. Again, (b) might just be Fred's (direct) awareness of the sun.

The preceding points, (1)–(4), collectively suggest that being "directly aware" of mental states amounts merely to these states being *conscious* states. So construed, talk of direct "awareness of" these states needn't mean they are themselves objects of cognition in the same way in which external objects are. That Fred is *consciously* conceiving the sun, in other words, doesn't imply that his conceiving is indirect. This point in turn raises an interesting possibility: If the qualitative character of a cognitive state may be linked with its being a conscious state, then Fred's direct awareness even of that character need not mean that it is itself an object of cognition in the same way external objects may be. But if so, then even his being sensorily aware (*prima facie*) of yellowness needn't preclude his also being aware of (say) some mechanical property (more in Section 8).

(5) Finally, what drives (A) is the non-identity between

(t) The state of cognizing the sun
(s) The sun itself

---

[21] *2nd Replies*, AT VII.160, CSM II. 113.

[22] *3rd Med*., AT VII.37, CSM II.25-6.

[23] Costa (1983): "… [T]he 'form' of my thought is that which makes it a thinking about Vienna rather than some other thought; it is simply that which all thoughts *of that type* have in common." (p. 542)

[24] *Med*., Preface, AT VII.8, CSM II.7.

which in turn suggests that awareness of the former cannot be identified with awareness of the latter. The most straightforward way to counter (A), therefore, would be to find some way to equate or identify (t) and (s). I stress "some way," for of course, (s) and (t) are not, as Descartes would say, *really* identical.[25] Nevertheless we clearly find in Descartes evidence of something like this strategy.

We see one hint in the distinction between ways of construing ideas. In (4) we took that distinction perhaps as a semantic one: "Idea" has two denotations, viz. mental acts and the (perhaps physical) objects of those acts. On that reading, mental acts are really distinct from their objects, and the word "idea" can be used to refer to either. In contrast, suppose that distinction were a *metaphysical* one: ideas are (single) entities either with a twofold nature, or whose nature can be grasped in two ways. On this reading, an idea just is an entity which is both (qua mental state) intrinsically related to us but also (qua representation) intrinsically related to the object represented.[26] The latter relation requires illumination, of course. But for now the key point again is this: given this twofold nature of ideas, non-reflective awareness of our mental state just is a form of direct awareness of the object. The twofold nature amounts, I suggest, to Descartes' seeing (s) and (t) as *in some sense* identical. Fred's thought of the sun just *is* the sun, in *some* sense.

That he takes this strategy, at least with the intellect, is most evident in his famous invocation of objective being.[27]

> 'Objective being in the intellect' … will signify the object's being in the intellect in the way in which its objects are normally there. By this I mean that the idea of the sun is the sun itself existing in the intellect—not of course formally existing, as it does in the heavens, but objectively existing, i.e., in the way in which objects normally are in the intellect.[28]

This notion has a long scholastic history, and its precise meaning both throughout its history and in Descartes is controversial,[29] but one way to understand it is right on the surface. In *1st Replies* Descartes makes clear that objective being neither is an "extraneous" label "which adds nothing to the thing itself,"[30] nor merely refers

---

[25] Descartes distinguishes between real, modal, and conceptual distinctions in *Principles* I. 60ff., AT VIIA.28ff., CSM I.213ff.

[26] In the *Meditations* Preface Descartes explicitly has the semantic distinction in mind (*Med.*, Preface, AT VII.8, CSM II.7), but in the *4th Replies,* where his terminology changes but not the substance of the distinction, he seems to be thinking metaphysically: AT VII.232, CSM II.163. Cf. Arnauld's *True/False*, Ch. 5, 66, *VFI* 38.198; and *3rd Med.*, AT VII.42, CSM II.29: ideas, while being mental modes, are also just the kind of thing which "contain" objects.

[27] I'll ignore possibly subtle distinctions between objective being and objective reality (Nelson 1996), and treat the following locutions as synonymous: "object x has objective being (in the intellect or mind)," "x exists objectively," "the idea of x contains or manifests the objective being of x," "the idea of x is the objective (mode of) being of x." Some texts: *3rd Med.*, AT VII.40, CSM II.28; *3rd Med.*, AT VII.41, CSM II.29; *2nd Replies*, AT VII.161, CSM II.113.

[28] *1st Replies*, AT VII.102, CSM II.75; cf. *2nd Replies*, AT VII.161, CSM II.113–114.

[29] Cf. Cronin (1966), Wells (1967, 1990), O'Neill (1974), Normore (1986).

[30] AT VII.102, CSM II.74.

to the fact that a mental act is a contemplating of some thing; rather it is some (intrinsic) feature *of the thing itself*. The "sun itself" exists formally, in the heavens, then, but the *very same* object can also exist objectively, in an intellect. Make no mistake: Descartes explicitly uses the phrase "the sun itself" to refer to *both* the formally and objectively existing sun. Fred's thinking of the sun, then, is not entirely external to the sun, but in some way involves *the very same sun* which exists "out there." Although this is an alien-sounding thesis to our contemporary ears, there are good reasons to ascribe it to Descartes: not least of these is his quite explicit statement of it just noted.[31]

How exactly to *implement* this thesis is not obvious. Cartesian dualism, with its stress on the distinctness of the mental and the material, for example, would seem resistant to it.[32] The thesis might require something like the scholastic forms which Descartes is generally taken to reject, viz. forms capable both of informing matter and existing in an immaterial mind.[33] Still, while Descartes' precise attitude towards "forms" is not easy to determine,[34] we might soften the blow by invoking

[31] There are several reasons to accept this strong conception of objective being rather than (a) that which, as Yolton (1984) puts it, "is most usually assigned by commentators" to Descartes (p. 38), viz. that for x to be in the understanding is for some representation distinct from x to exist in the mind, or (b) that which Yolton (1984) himself defends, according to which it is merely for x to be understood (pp. 38–39): (i) The strong conception best fits his actual language. (ii) It seems required for his Third Meditation proof of the existence of God. That proof invokes the contentious principle that the total cause of an idea must ultimately have as much formal reality as is contained objectively in the idea itself (*3rd Med.*, AT VII.41, CSM II.28–29), a principle derived from the more general one that there is nothing in the effect which didn't previously exist in the total cause. The strong conception of objective being perhaps explains both the contentious principle and precisely why Descartes derives it from the more general one: The idea of the sun quite literally involves "the sun itself," so the cause of the idea must be comparable to whatever could cause the sun itself. As Aquila (1974–5) writes: "But Descartes could not have supposed his proof to have any force at all unless he had already been prepared to suppose that an infinitely perfect nature does at least exist *in* the understanding in something more than the sense that there exist mental acts by which such a nature is contemplated." (p. 234) (Cf. O'Neil 197, p. 72.) (iii) Aquila (1974–5) argues that the very possibility of non-reflective awareness of the sort required by Cartesian transparency requires that an idea's being directed towards some object must not be distinct from some object's coming to exist in the mind (p. 237).

[32] Cf. Clemenson (2005).

[33] Cf. Lennon (1974). For scholastic theories of cognition, see Pasnau (1997).

[34] Sometimes he's explicitly critical of them (*Principles* IV.198, AT VIIIA.322, CSM I.285; To De Launay, 7/22/41, AT III.420, CSM III.188); sometimes he deliberately just ignores them (*Meteorology* I; AT VI.239, To Morin, 7/13/38, AT II.200, CSM III.107); yet sometimes he invokes them (*5th Replies*, AT VII.356, CSM II.246; *2nd Replies*, AT VII.160, CSM II.113). Nevertheless his account of objective being seems heavily indebted to the Aristotelian-Scholastic one, even to the point where his "simple natures" or "essences" seem equivalent to the substantial forms he supposedly rejects. (See O'Neil (1974, Ch. 1, pp. 73–74) for texts from Aquinas and Aristotle illustrating that "Descartes' overall position on *esse objectivum* is in keeping with the Tradition." (p. 73)) In general, as I argue in Pessin (2007), he is best construed not as rejecting scholastic forms but as reconceiving them: *Optics* AT VI.112, CSM I.165; *Principles* IV.198, AT VIIIA.322–23, CSM I.285.

more Cartesian terms, viz. "essences" or "simple natures." These Descartes held both to be grasped "directly" by the intellect and ultimately to constitute the physical world.[35] To say the sun itself exists in the intellect would, then, be to say that a certain essence or nature, of a body with certain mechanical properties including size and location, exists objectively, the very same essence which, realized in nature, is the formal sun.[36]

Fortunately, this should be sufficient for our current purposes. Fred's act of thinking of the sun is just the sun itself's way of being in his intellect, as just described; his thinking of the sun is therefore identical, in that sense, with the sun itself. But then (s) and (t) are in that sense identical; but then Fred's non-reflective awareness of his thinking of the sun, via transparency, just *is*, in that same sense, his direct awareness of the sun itself. There is no veil.

Now other scholars have applied Descartes' conception of objective being in defense of the directness of Cartesian intellectual cognition.[37] It's fairly clear how such a defense might go, as sketched above, once we accept that the same (mechanical) essences can be realized both in matter and in intellect; indeed we essentially just have a version of the standard scholastic account of direct realism, with mechanical essences substituted for forms. It is much harder to see how a similar account could be given of the directness of sensory cognition, however, in light of Descartes' general exclusion of secondary qualities from the physical world. Nevertheless that is just what I aim to do.

3 The Proposal (P)

Here is my proposal, (P): Just as the Cartesian thought of the sun is the sun itself in the intellect in the way objects are "normally there," so too the sensation of yellow is motion-m in the senses in the way objects (or properties) are normally there.[38] Just as Fred's thought of the sun just *is* the sun, thought, so too sensation-y

[35] Cf. O'Neil (1974, Ch. 1).

[36] I develop this account in more detail in Pessin (2007). Implicit in it is the important point that, while (say) the formally existing sun is really distinct from the objectively existing sun, the "essence" realized in matter could not be said to be really distinct from the essence realized in the intellect. Despite the natural tendency to read Descartes' treatment of universals (for example in *Principles* I.58ff., AT VIIIA.27ff., CSM I.212ff.) as nominalistic, then, I argue that we must read him as subscribing to a realism.

[37] O'Neil (1974, pp. 71ff., 88ff.), Aquila (1974–5, p. 236), Lennon (1974), MacKenzie (1990), Chappell (1986), and Hoffman (2002) reject this strategy.

[38] Objects needn't be "in" the senses in some way exactly reflecting their extra-mental being. Here we might see Descartes respecting the well-known Scholastic axiom: "The thing known is in the mind of the knower after the fashion of the knower" (cf. *Rules* 12, AT X.423, CSM I.47; O'Neill (1974, pp. 31, 73) provides texts from Aquinas and Aristotle illustrating the axiom). In general, Descartes' account of the sensory process is highly sensitive to the degree to which the subject's own properties and receptivity influence the sensation produced (e.g., *Optics* 6, AT

just *is* motion-m, sensed.[39] Just as Fred's direct (i.e., non-reflective) awareness of his thinking of the sun just *is* his (cognitively) direct awareness of the sun, so, too, Fred's direct (non-reflective) awareness of his sensing of yellow just is his (cognitively) direct awareness of motion-m. In sensing the yellow, Fred in fact is sensing motion-m (despite his perhaps not realizing this). If so, as I shall argue, then Cartesian sensation is also direct.[40]

4 Sensation, Objective Being, and Representation

We might begin by asking whether Cartesian sensations contain objective being as intellections do. There are reasons to think not;[41] but then there are reasons to think so.[42] To make some headway, consider that "objective being" is at least part

VI.130–147, CSM I.167–175). The very same physical property—motion-m—may thus be "in" the different cognitive faculties, and provide the content of our states, in very different (qualitative) ways: *World* 2, AT XI.9, CSM I.84; *3rd Med.*, AT VII.81, CSM II.56; To Regius, January 1642, AT III.493, CSM III.206. (See Section 9 for further discussion.)

[39] The theory must accommodate the actual etiology of sensations, of course, which always occur in particular sensory or causal contexts: just which states of the physical world result in just which pineal gland states result in just which sensory states is a complex matter determined by the laws of physics and the body–mind laws (and ultimately by God's benevolence). The very same motion-m might ultimately result, for example, in different pineal gland states, hence different sensory states, as in the *World* text of the previous note; and perhaps different motions ultimately might, in different contexts, result in the same pineal gland state, hence same sensory state (as in the dropsy case of the *6th Med*.). Nevertheless, these facts needn't entail that the representational content of sensory states itself reflects contextual or relational physical properties; God might well arrange that the qualitative character of sensations does this, while their representational content remains the relevant motion-m. (Again, see Section 9 for more.)

[40] Hoffman (1996) suggests something like (P) in passing (p. 378), but resists the "direct cognition" conclusion in Hoffman (2002).

[41] (i) He invokes "objective being" mainly with respect to the intellect. (*1st Replies*, AT VII.102, CSM II.75; *Med*. Preface, AT VII.8, CSM II.7; *3rd Med.*, AT VII.41, CSM II.29.) (ii) In one of the proofs of God's existence, Descartes seems concerned to restrict his discussion to the objective being of ideas he grasps clearly and distinctly (*3rd Med.*, AT VII.43–44, CSM II.29–30). (iii) Descartes (a) writes with respect to paradigm sensory ideas that "the reality which they represent [*exhibent*] is so extremely slight that I cannot even distinguish it from a non-thing," (*3rd Med.*, AT VII.44, CSM II.30) and (b) seems to agree with Arnauld that if cold is a privation, "the idea of cold is not coldness itself as it exists objectively in the intellect," (*4th Replies*, AT VII.233, CSM II.163) in both cases perhaps suggesting that sensory ideas lack objective being.

[42] (i) Some texts suggest a broader applicability beyond the intellect (*2nd Replies*, AT VII.161, CSM II.113–114; *3rd Med.*, AT VII.42, CSM II.29). Descartes often speaks of ideas of sensation, including shortly after noting that ideas contain objective being "by their very nature" (*3rd Med.*, AT VII.43, CSM II.30). If the objective being contained in such an idea is "the being of the thing which is represented" by the idea, or if "whatever we perceive" is "whatever has objective being in one of our ideas" then there seems no reason not to think that our sensory ideas themselves have objective reality. (ii) Those texts which do restrict themselves to intellect might, perhaps, be using "intellect" loosely to refer to the mind in general. (Cf. Simmons (2003,

of a theory of mental representation:[43] that a mental state contains objective being, that something exists objectively therein, and that the state represents the thing, are intimately linked. Where there is objective being, at least, there is representation. The crucial question is whether the converse obtains.

But *that* question is plagued by an important complication. Given its central role in his epistemology, any Cartesian account of mental representation must be sensitive to epistemic concerns. Since sensation and intellection play such diverse epistemic roles, we might expect differences in their representational nature. It may be precisely because the Cartesian intellect provides us with a "better" kind or greater degree of knowledge of the world than do the senses, for example, that "objective being" finds its *primary* or clearest application with respect to the intellect.[44] But that leaves two options: (i) allow that sensations, if they are representations, contain objective being despite their epistemic limits; or (ii) deny they do, despite being representations, because of their limits. On (i) we would say that where there is representation there is objective being, but on (ii) we would not.

Though I believe we should go with (i),[45] in fact not much rides on it for our purposes. What (P) needs is merely that Cartesian intellection and sensation should be analogous in relevant respects; and that analogy requires so far only that sensations and intellections both represent the external world. Their manner of representation may differ, epistemically, etiologically, and even teleologically.[46]

---

pp. 563–564, p. 569).) (iii) Generally, Descartes suggests that all ideas have representational content (*3rd Med.*, AT VII.44, CSM II.30) and often refers to sensory ideas (*3rd Med.*, AT VII.43, CSM II.30); if containing objective being is a matter of having representational content, then sensory ideas seem as good as intellectual ones. (iv) Nor does he always restrict objective being to ideas he grasps clearly and distinctly: *3rd Med.*, AT VI.41, CSM II.28; *1st Replies*, AT VII.103–104, CSM II.75.

[43] Descartes roughly characterizes objective reality in terms of representation: *3rd Med.*, AT VII.83, CSM II.28.

[44] For example, ideas grasped clearly and distinctly are such that we can determine the true nature of what we cognize, given the truth rule and the consequent correspondence between objective being and formal being. But in sensing yellow, we cannot determine its true nature, i.e., *what* we are cognizing.

[45] Both for the reasons sketched in note 42 and because the opposing reasons (from note 41) can, I think, be countered. In *3rd Med.*, for example, Descartes is primarily troubled by his inability to *tell* exactly how much reality (or what being) is contained in his sensations; this needn't mean that in his final, considered view they lack objective reality, and indeed he admits that as long as they do *not* derive from a privation (which, ultimately, they do not) then they *do* contain some ("extremely slight") reality thus being. Moreover, at this stage of the Meditations he is limiting himself to affirming only what is certain, and since the objective content of sensations is obscure he must here, in effect, reject it. It's also no surprise that in his proof for God's existence he restricts himself to ideas he grasps clearly and distinctly, for these are the ideas about whose objective being he is most certain. That itself is irrelevant to whether sensory ideas might themselves contain objective being.

[46] So, for example, sensations are epistemically opaque, mediated by the agent's body, and aimed at the preservation of the mind–body union all in a way intellections aren't. (See Sections 5.2 and 9 for further discussion.)

There may well be epistemic reasons to limit objective being to the intellect. But all this means is that we ought to distinguish between sensations being representational and our being able to determine that (and what) they represent, in a way we don't need to for intellections. The analogy we seek does not require comparable epistemic access to the objects of our cognitive states; it merely requires that there *be* some object (whether or not we can tell).

So we might next ask whether Cartesian sensations represent external objects or properties. Simmons summarizes three strands of thought, each of which receives textual support:

(a) On the first strand, sensations do not represent, or even purport to; instead they are merely mental modes, or the mind's awareness thereof, lacking reference to extra-mental reality.[47]

(b) On the second, sensations do purport to represent, but simply fail to: nothing extra-mental properly corresponds to them.[48]

(c) On the third, sensations actually succeed in representing the extra-mental, with the catch being that what they succeed in representing may not be exactly as is apparent to the senses. Simmons writes: "The thought here is that my sensory experience represents the corporeal world to me, but I may not be able to tell, on the basis of that experience, its true nature."[49]

Since the question of whether Cartesian sensations are representational is hotly debated,[50] space considerations shall lead me merely to explain briefly why I think (c) is preferable, without offering full argument. (If my thesis must consequently be restricted to the claim that Cartesian sensations represent directly *if* they represent, then so be it.)

(1) I find (b) implausible. What underwrites the claim that sensations fail to correspond to anything external is their non-resemblance to mechanical properties; but, as we'll see below, that failure is perfectly consistent with their representing those properties. Moreover, intellectual ideas are paradigm representational states, even where nothing external exists to correspond to them, suggesting that successful representation need not require a state's having external relations. More generally, the very distinction between "purporting" and "succeeding" to represent is quite problematic: its notion of "purport" and that notion's relationship to judgement are not clearly worked out, it also requires (as we'll see) a representing state to have external relations, and it seems incapable of ensuring that the *representandum*

---

[47] *Principles* I.71, AT VIIIA.35, CSM I.218–219, and *Principles* I.68, AT VIII.33, CSM II.217; also *6th Replies*, AT VII.440, CSM II.297; To Chanut, 2/26/49, AT V.292, CSM III.369. MacKenzie (1990) supports this reading (p. 123).

[48] *3rd Med.*, AT VII.37, CSM II.25; AT VII.44, CSM II.30.

[49] Simmons (1999, p. 350). Simmons cites texts such as *6th Med.*, AT VII.80ff., CSM II.55ff., To Hyperaspistes, August 1641, AT III.424, CSM III.189; To More, 2/5/49, AT V.271, CSM III.362. Cf. *Principles* I.70, AT VIIIA.34, CSM I.218. Simmons ascribes this interpretation to Alanen (1994), Bolton (1986), Schmaltz (1992), and Wilson (1990), as well as to Arnauld.

[50] Cf. Bolton (1986), Wilson (1990), Schmaltz (1992), Hoffman (1996), Simmons (1999), etc.

genuinely count as an object "of" (or "in") thought, and so fits poorly with much of our discussion in Section 2.

(2) I'm persuaded by the cases made by various of (c)'s proponents.[51]

(3) As we'll explore below, Descartes strongly stresses the unified nature of the mind and its distinction from the body. This nature is best served, I think, by treating all mental states, and no physical states, as being inherently representational states.

(4) Finally, given that Cartesian sensations (i) *are* mental states, possessed of qualitative characters which neither (ii) "resemble" the relevant mechanical properties (as we'll discuss) nor (iii) serve well as epistemic resources, I can understand why Descartes might sometimes, in texts invoked for (a), speak of them as if they were non-representational. Yet (I believe) (i)–(iii) are all perfectly consistent with (c).

I'll henceforth take Cartesian sensations to be representational. That's a start; but (P) requires us to go a bit further into the nature of that representation.

Now, sensations do bear various relations to extra-mental reality. They are brought about, for example, in generally nomic ways, proximally by bodily states and distally by more remote states of matter.[52] If they fail to resemble the external world, they in some sense "correspond" adequately to it to help the embodied mind navigate the world successfully. But are these sorts of relations sufficient to *ground* the representationality of sensations?

Suppose, for a moment, that sensation-y represented motion-m by virtue of some external relation between y and m. If so, then Fred's awareness of his state of sensing would be very hard to identify with his awareness of m, since nothing about that state itself (or his awareness thereof) would contain or indicate m. Indeed, if nothing intrinsic to his sensing yellow indicates m, Fred should not be said to be sensing m at all. (P) requires, then, that the representational property be an intrinsic or internal one: something about sensation-y *intrinsically* makes it a representation of motion-m. Such a requirement has a good precedent, of course: as we saw, Descartes considered objective being, at the heart of Cartesian representation, both to be an "intrinsic denomination" of the object represented and due to the "very nature" of ideas.[53] These strongly suggest an internal relation between the idea and its *representandum*. We ought not to demand anything different from sensory ideas.[54]

---

[51] For example, by: (i) Wilson's (1990) general account of the representational nature of all Cartesian mental states; (ii) Simmons's (1999) case that the bio-functional role Descartes assigns to sensations requires that they represent. (iii) Aquila's (1995) case that Cartesian sensations themselves "contain" (thus represent) the "teaching of nature" that we are united to a body. (Cf. *6th Med.*, AT VII.81–82, CSM II.56; *Principles* II.3 AT VIIIA.41–42, CSM I.224.)

[52] There is dispute about whether there is a genuinely causal relationship between brain and mind, but not that there is a generally nomic relationship between brain states and mental states. Cf. Pessin (2003) for discussion.

[53] *3rd Med.*, AT VII.42, CSM II.29.

[54] One might object by noting that with a sensory idea, unlike an intellectual one, its obscurity prevents our determining by reflection much about its *representandum*, and that an explanation

Further, sensory representation might well, as we'll see, simply be a primitive for Descartes. If so, then there may be nothing about a sensation to which we could point and say "that is the feature by virtue of which it represents." That might help explain our inability to determine the true nature of the *representandum* of our sensory ideas: that which *is* in this way accessible to us (either by transparency or by reflection) might just not be sufficient to "explain" the representing. Thus Fred's sensation-y might intrinsically represent motion-m, be an idea *of* m, despite his being unable to determine this.

With these points in mind, let's consider two recent theories of Cartesian sensory representation.

5 Theories of Sensory Representation

5.1 A Causal Theory of Sensory Representation

"The perceptions we refer to things outside us," Descartes writes, "namely to the objects of our senses, are caused by these objects, at least when our judgements are not false … And we refer these sensations to the subjects we suppose to be their causes …."[55] Texts such as this one, combined with Descartes' detailed analysis of the etiology of sensation, may suggest a causal theory of sensory representation. The basic idea would be that a sensory state represents that which causes it "in the right way." After distinguishing two notions of representational content in Descartes, Wilson tentatively proposes, for one of them, just that: for an idea or mental state "referentially to represent a certain physical state is just for that idea to be caused—in the 'right' way—by that state, whatever it might be."[56]

Wilson herself, however, has reservations about the theory: "On the whole I suspect that the causal account was influential in Descartes' thought, even if he was unable to develop it fully, to create a theory immune to counter-examples."[57] One counterexample she has mind is that Descartes thinks that even non-existing objects can be referentially represented, despite their inability to play the requisite

---

of this difference might be that intellectual representation is internal to the idea in a way that sensory representation is not. Perhaps; but this epistemic difference does not require that explanation. As noted above, there are important etiological and teleological differences between sensory and intellectual cognition; and in any case the demands of Cartesian "knowledge" are quite complex. We'd need much further argument that these facts fail to explain the epistemic asymmetries before we should assign the differences to the natures of sensory and intellectual representation themselves. In any case, the unity of the Cartesian mind and its distinction from the body as mentioned above strongly suggest a unified account of representation, thus an internal sensory representational relation as well.

[55] *Passions* I.23, AT XI.346, CSM I.337.

[56] Wilson (1990, p. 75). Cf. Normore (1986), MacKenzie (1990), and Nelson (1996).

[57] Wilson (1990, p. 76).

role in a causal theory. Simmons (1999) similarly objects that a simple causal theory would fail to distinguish between the proximal (say, pineal gland) cause of a particular state and any of its many more distal causes. While I'm not persuaded by these particular problems, there are others:

(1) Wilson distinguishes between "presentational" representational content, roughly "what the mind takes itself to be aware of,"[58] and "referential" representational content, roughly what the mental state is ultimately referred to (even if the mind is unaware of it). She proposes the causal theory to account only for the latter. At best, then, the theory leaves unexplained what is arguably the more basic notion of Cartesian representation.

(2) More generally, the "presentational"-"referential" distinction is problematic in ways similar to those troubling position (b) just above. (Indeed having "presentational" content might be one way in which sensations "purport" to represent.) In particular, nothing ensures that the object "referred" to is present "to" or "in" thought in the right way, which suggests that referential content is not really genuine Cartesian mental representation. Or put differently: the distinction fails to ensure that the relevant element in the causal chain is an object of *cognition*. Moreover, presentational content is poorly defined: 'what the mind takes itself to be aware of" perhaps allows or overlooks a gap between what the mind is *in fact* aware of and what it "takes itself" to be aware of, where the latter might involve the intellect, or even the will, in ways inadequately specified. What really matters for our concerns, in any case, is what the sensing mind is *in fact* aware of, whether or not it successfully "takes itself" to be aware of it (more in Section 7).

(3) Descartes nowhere explicitly suggests a causal theory of content despite his extensive work on the causal processes producing sensation.[59]

(4) Causal relation in general is not sufficient for representation. Simmons (1999) points out that in the physical world an effect (such as the melting of wax) may be thought of as a (reliable) sign or indicator of its cause (such as heat), but is not typically thought of as a representation of it. What, she asks, is different about the sensation of heat? In fact I think this point can be put more forcefully: Cartesian dualism requires a significant ontological difference between mind and body. As we'll discuss in Section 5.3, the ability to represent would provide just that; the ability to "be an effect" (or a "sign") does not. But then the causal relation between x and y isn't sufficient to make y represent x unless y is *already* the sort of thing capable of representing.

(5) More importantly, causal relation is not necessary for representation *simpliciter*. The paradigm example of Cartesian representation is, again, intellectual representation:

---

[58] *Ibid*., p. 81, n. 10.

[59] Even the *Passions* text cited to start this section actually takes the representational content for granted: it's only because we already identify the sensation as "of" a torch or "of" a bell that we know to refer the sensation to those objects in the course of making judgements. Similarly, when Descartes explores the causal origin of sensations in the *6th Med*. the issue is whether sensations are caused by bodies, hence support "true" judgements about the world, and not whether or what they represent.

(i) But the connection between intellectual ideas and their "causes" is far looser than that between sensory ideas and their causes. Our thinking of the sun is not typically caused by any particular circumstances, physical or mental. Indeed the general independence of the intellect from the body/brain, according to Descartes,[60] tells against assigning intellectual ideas any physical causes, which would undermine analyzing their content in terms of their causes.

(ii) In addition, we enjoy some kind of first-person privilege about the contents of our intellectual states. That means that in some sense we are infallibly aware of what we're thinking of, whether or not it exists. But we are not aware in that sense of the causal relations our intellectual states hold to external objects, allegedly constituting their content.

(iii) In Section 2 we saw reason to hold that the representational nature of intellectual states is intrinsic to them. Their external causal relations would simply be irrelevant to their representing.

In general, then, causal relations seem irrelevant to Cartesian mental representation. That's perfectly consistent with admitting that the causal relations between sensations and the world are very important; it just means we should not expect sensations to represent by virtue of those relations. If anything the non-deceiving Cartesian God sets up the causal relations He does between the physical and the mental because the mental states represent what they do; representation *precedes* causation.

But even more generally: Wilson's very distinction between "presentational" and "referential" content might already presuppose the veil of perception, by distinguishing what we we're aware of (or "take ourselves" to be aware of) from what is the ultimate referent (or object) of our perception. But what much of the preceding critique suggests is that the notion of "referential" content simply is not one of genuine Cartesian mental representation. Causal relations may (perhaps) establish various correspondences between mental state and external objects, but that in itself doesn't establish that those objects are the objects of cognition. "Presentational" content is not merely the more basic notion of representation, then, but really—once its kinks are worked out, below—the *only* notion of Cartesian representation. And Wilson's theory is silent on the question of what *its* objects are, and thus on the question of the veil.

5.2 A Bio-Functional Theory of Sensory Representation

Simmons writes:

> …[T]he function of the senses [is] biological or ecological: they enable the mind-body union to survive in its physical environment. It is also, however, a cognitive or representational function. Sensations conduce to self-preservation by *showing the mind*

[60] To Reneri for Pollot, April or May 1638, AT II.38, CSM III.99.

> *what bodies (its own included) are like*, not in themselves as conceived by the Cartesian physicist, but *relative to its own body's well-being*. In other words, they represent to the mind ecologically salient properties of (or perhaps facts about) the corporeal world: Where are external bodies relative to where my body is right now? Will they pose a threat to my body? Will they promote its health and fitness? Is my body damaged? Is it healthy? …. [S]ensations … represent ecological properties of the corporeal world: pains represent bodily damage, tickles represent bodily health, foul taste sensations represent the undigestibility of bodies, colour sensations represent surface differences, and so on …. My suggestion, then, is that Cartesian sensations represent things in the corporeal world … in virtue of the role that they play in enabling us to interact with the world in a self-preserving way.[61]

There seem to be two distinct theses in play here:

(a) What grounds Cartesian sensory states' having representational content is their "bio-functional" role.
(b) The content of those states is "ecological": *what* they represent are ecological (or ecologically salient) properties such as bodily damage or health, (un)digestibility of food, etc.

Most of Simmons' paper develops and defends (b), but her suggestion that sensations represent "in virtue of their role" strongly implies (a), so I'll consider both. Further, she offers her theses as an alternative to Wilson's causal theory specifically of referential content; like the causal theory, then, the bio-functional theory is an incomplete theory of Cartesian mental representation, since it offers no account of presentational content. My critique aims to show that the bio-functional theory fails, as did the causal theory, as a theory of Cartesian mental representation *simpliciter*; since both are theories of referential content, much of what I say below amounts to further argument that referential content just is not Cartesian mental representation.

5.2.1 Does Bio-Functional Role Ground Representational Content?

We may interpret this suggestion two ways. First, the "bio-functional role" of a sensation is what its job is, what it is designed or intended or supposed to do.[62] Pain (for example) might then represent bodily damage because pain has the role of alerting the mind to the damage etc. So interpreted, there's nothing objectionable here, but also nothing informative: we have learned to what end a pain is made representational, but not what actually *makes* it representational, and so we lack a theory of representation. Indeed the natural question here would be why God would give pain that particular role, and not some other sensation, if not because pain, in *already* representing damage, was so well-suited to it?

---

[61] Simmons (1999, pp. 355–357).

[62] That Descartes, opposed to invoking "final causes" in physics, may be read as invoking bio-functional roles so construed is defended by Simmons (2001).

On a second, more interesting, interpretation, the bio-functional role of a sensation is its location in a particular causal network. Sensations have representational content by virtue of having particular sorts of causes *and* their causing (or disposing) the mind to act in certain "self-preserving" ways. Functional roles, or ecologically-oriented causal connections, in other words, *bestow* content on sensations. Interpreted this way, (a) is indeed a "theory of representation."

Still, I don't believe it works for Descartes.

(1) First, having a "functional role" is not sufficient for being a representation. The physical world is full of states which are not considered representations despite having causes and effects. Even having a specifically "bio-functional" role is insufficient: Cartesian animals behave in self-preserving ways but lack mental properties including sensations,[63] and thus have physical states with bio-functional roles, which, nevertheless, are not representations.

(2) Similarly, functional role is generally not relevant to representational content.

(i) Suppose Fred has a painful sensation "in" his foot. That this sensation is "located" in the foot is an inseparable aspect of it; Fred doesn't feel a generalized pain, but a pain-in-the-foot.[64] This representational content appears to be something about which Fred has a first-person privilege. But it's implausible to hold that he is in that sense aware of the complete causal ancestry allegedly constituting its content. And even if he is aware of the most salient element of the causal sequence—viz. the foot—he is not aware of the foot qua cause of the sensation. In fact just the reverse: he considers the foot the (distal) cause of his sensation by virtue of the sensation's representing his foot. But then representing precedes the functional role.

(ii) In general, the theory gets things backwards. It's because a state represents what it does that it's well-suited to play the functional role it does; that a pain sensation represents bodily damage (say) would explain why God establishes that it be caused by that damage and tend to cause relevant responses. But then, again, representing *precedes* functional role. Otherwise, any arbitrary state that God might nomically link to bodily damage and avoidance behavior would, by definition, have counted as a pain sensation. As Simmons herself points out, for Descartes, God's choice of which pineal gland states to map with which sensory states, via body–mind laws, is *not* entirely arbitrary:

> It is made with regard to the ability of sensation-types to alert the mind to aspects of the corporeal world that are relevant to the continued survival of its body (and so of its union with the body). Nature (God) has paired each type of pineal motion with that type of sensation which 'of all possible sensations, is the most especially and most frequently conducive to the preservation of the healthy person'.[65]

---

[63] To Gibieuf, 1/19/42, AT III.479, CSM III.203–204; To More, 2/5/49, AT V.277–279, CSM III.366.

[64] *6th Med*., AT VII.88, CSM II.60.

[65] Simmons (1999, p. 357), citing *6th Med.*, AT VII.87.

Why would Descartes need to explain why God chose just *this* sensation if *any* sensation causally linked in the right way would have served as well?

(iii) There is, however, another sense in which God's choice *is* arbitrary: there is nothing which *intrinsically* or necessarily relates a given pineal gland motion with any particular sensory state.[66] God could have willed the body–mind laws differently; had He done so, the same pineal gland motion could have caused a different sensory state *with a different content*, or the same sensory state with its same content could have been caused by a different pineal gland motion. But then a sensory state's causal ancestry—part of its functional role—is irrelevant to its content. God's benevolence ensures that He not will the laws in this arbitrary way; but on the bio-functional theory, logic itself would forbid it. The conclusion is the same: representing precedes functional role.

(iv) The preceding paragraph assumed that the sensation's effects remained constant. But perhaps representational content is determined (partly) by the *effects* of the sensory state in question? Does a sensation of thirst, for example, represent the body's need to drink by virtue of its causing, or being, a disposition to drink,[67] or is that relationship itself also "arbitrary" in the sense just described? Descartes writes: "…[I]n thirst the sensation of the dryness of the throat is a confused thought which disposes the soul to desire to drink, but is not identical with that desire."[68] The distinction here between sensation and desire implies a contingent relationship between them, such that God *could have* causally joined that same sensory state with some other effect without thereby changing the content of the sensation, or perhaps joined some other sensory state with the disposition to desire to drink without turning that other sensation into a representation of the body's need for water.[69] From theodicy concerns, of course, God *wouldn't* have; but the fact that He *could* have is enough to deny that their causal consequences determines their content.[70]

(v) Similarly, although Descartes' account of free will is murky, if you hold that he generally subscribes to a libertarian view (outside of assent to clear and

---

[66] *6th Med.*, AT VII.88, CSM II.60–61; *6th Med.*, AT VII.76, CSM II.52–53. Cf. Cottingham (1986, p. 139ff.); Wilson (1991, p. 43).

[67] *6th Med.*, AT VII.88, CSM II.61.

[68] To Chanut, 2/1/47, AT IV.603, CSM III.306–307.

[69] Cf. *Principles* IV.190, AT VIIIA.317–318, CSM I.281. While Simmons (1999) suggests there is something "intrinsic to the sensation of pain" which elicits the appropriate self-preserving response (p. 358), she also inclines towards the idea that some Cartesian sensations "phenomenologically present their objects in a less evaluative, simply descriptive, way and then give rise, naturally and automatically (according to the institution of nature), to a pursuit or avoidance response." (p. 366, n. 20) In support of this possibility she notes *6th Med.*, AT VII.82, CSM II.57. This and other texts similarly suggest that the avoidance/pursuit responses are importantly distinguishable from the preceding sensations, and that God could have joined different responses to these sensations. (See also *6th Med.*, AT VII.76, CSM II.52–53.)

[70] If this is so for internal sensations such as thirst, then all the more so with respect to external sensations such as Fred's sensation-y, which are not correlated with any sort of typical effects.

distinct propositions), then we can freely choose how to respond to any of our sensory states. If so, there could be no definitional relationship between sensory representational content and the effects of the sensory state.

(3) Finally, having a particular functional role is just not necessary for having representational content *simpliciter*. The functional role theory is as ill-suited to intellectual ideas as was the causal theory, since their causal connections are far looser than those of sensory ideas. Indeed, that Cartesian intellectual ideas differ in function from sensory—in aiming towards truth—shows that ideas can represent without any particular bio-functional role. Finally, again, our first-person privilege with respect to intellectual states means that we are infallibly aware of what we're thinking of. But it's implausible to hold that we are aware in that sense of the functional roles allegedly constituting their content. That, again, implies that the representational nature of those states is independent of bio-functional role.

5.2.2 Do Ecological Properties Constitute Sensory Representational Content?

Granting that the "ecological function" of sensation is, as Simmons puts it, "to guide our successful, i.e., self-preserving, interaction with bodies in [the] local environment,"[71] the ecological properties of bodies might be those properties cognition of which is most conducive to our successful interaction. Although broader notions are available, most of Simmons' article indicates that she takes these to be such properties as being harmful or beneficial (in various ways) to us. But do these constitute the content of Cartesian sensations?

(1) The main problem with thesis (b) is that it isn't easily applicable to most sensory states. Its best case is the "internal sensations" of pain and pleasure, for these are well-suited to play an "ecological" role. But despite sometimes speaking more generally,[72] Descartes seems to ascribe an ecological function *specifically* to pain and pleasure alone. Speaking of external sensations, he writes: "…[T]he fact that some of the perceptions are agreeable to me while others are disagreeable makes it quite certain that my body … can be affected by the various beneficial or harmful bodies which surround it."[73] Descartes here maps the *(dis)agreeableness* of certain sensations with bodies' being harmful or beneficial, and notes that only "some" of his sensations display such qualities.[74] Since most external sensations

[71] Simmons (1999, p. 355).

[72] "For the proper purpose of the sensory perceptions given me by nature is simply to inform the mind of what is beneficial or harmful." (*6th Med.*, AT VII.83, CSM II.57) Of course, he may hold that this is the "proper purpose" of sensory perceptions without holding that every particular sensation fulfills this purpose.

[73] *6th Med.*, AT VII.81, CSM II.56. Cf. *6th Med.*, AT VII.82, CSM II.57; *Principles* IV.191, AT VIIIA.318, CSM I.282.

[74] He's even more explicit in *Principles* I.71, AT VIIIA.35, CSM I.218–219. See also *Principles* AT VIIIA.318, CSM IV.191; *6th Med*., AT VII.74–75, CSM II.52.

don't display a relevant form of "(dis)agreeableness," their contents are not good candidates for being ecological.

(2) Thesis (b) seems *false* for most external sensations. Colour differences, Simmons suggests, represent surface differences, so colours, presumably, represent surface textures. But while these may (though needn't) be ecologically *relevant* properties, sensory states don't seem to represent the ecologically relevant features of these textures. The harmfulness or beneficialness of certain bodies does not map in any regular way onto their surface textures. Not all yellow bodies, for example, are equally harmful or beneficial; not all harmful (or beneficial) bodies generate the same external sensations. Similarly, there is no general answer to how we do or should respond to a yellow body or how our responses to differently coloured bodies should vary. What yellow bodies do share is their surface texture; *what* our external sensory states represent, then, are not ecological properties but just the mechanical properties themselves.

In general, most external sensations are equivalent ecologically—say, equally pleasant or unpleasant—yet differ qualitatively. This, too, suggests, that what sensations represent are those features of objects with respect to which those objects differ, viz. their mechanical properties.

(3) Simmons writes:

> Now since these ecological properties must be instantiated in the corporeal world as modes of *res extensa*, we might say that sensations also represent modes of *res extensa*, but we should understand that they represent these things only indirectly, by way of representing the ecological properties they instantiate. What is important here is that the level at which sensations get their representational hook on the world is the level of ecology not physics.[75]

My final point is that even if sensations represent mechanical properties only "indirectly," in Simmons' sense, sensory cognition can be "direct" in the "non-veil" sense required for (P).[76] Consider several ways of analyzing Fred's sensing foot pain. We might say that

(i) Fred senses his foot *as* harmed, i.e., *as* damaged, and not *as* instantiating particular mechanical properties "motion-f." But then it is still *his foot* he is sensing. Or, perhaps

(ii) What Fred is sensing is the property of being damaged. But the foot possesses this property only insofar as it possesses motion-f. There is nothing more in the foot than its motion-f. If Fred senses the damage *in the foot*, it can only be by sensing motion-f in the foot; thus Fred senses this motion directly. Or alternatively still

(iii) Perhaps the property of being damaged involves a relation between motion-f and some other thing (such as the mind-body union), so in sensing the relation Fred may be said to be sensing both motion-f *plus* that other thing. Fine:

---

[75] Simmons (1999, p. 356).

[76] Simmons (1999) is actually sympathetic to this point, acknowledging that it is "probably correct" that Cartesian sensations are "obscure and confused perceptions of *res extensa*." (p. 356).

as long as motion-f is sensed, even along with other things, then it's sensed directly.

Although all three preserve direct sensation as required for (P), I believe (i) is the best route here. For either the "damage" is (roughly) identical to "motion-f," or "damage" is some relational property obtaining between (say) motion-f and (say) the mind–body union. Descartes is clear that sensations "do not always show us external bodies exactly as they are, but only in so far as they are related to us and can benefit or harm us[:]"[77] a sensation revealing "damage" qua ecological property would be revealing something relational. But motion-f is not relational in the relevant way, hence cannot be identified with damage. But then if "damage" is relational, it cannot be said to be instantiated or located in the foot (alone); so if, as Simmons suggests, pain represents "damage," it can't be said to represent something *in the foot (alone)*. Yet this seems wrong: what we experience is a pain-(*entirely*)-*in-the-foot*.[78] It also seems inconsistent with a point Simmons raises in support of her theory. With respect to "matters regarding the well-being of the body," Descartes notes, the senses "report the truth much more frequently than not."[79] But if the (internal) sense is reporting damage "(entirely) in the foot," it would be reporting a falsehood just where it should be reporting truth.

A better way: Deny altogether that "damage" constitutes part of the representational content of the sensation. Rather, *what* Fred senses is motion-f, just as option (i) suggests; but the mode or manner of that sensation, its qualitative character, reflects that motion's relational properties, or, more precisely, contributes to the sensation's biological function.[80] As the letter to More above suggests, sensations show us bodies insofar as they are related to us—but it is *bodies* they show us. Thus, roughly, Fred's senses *are* reporting the truth.[81] But of course all this is just what (P) says. Further, it's what (P) says, in effect, about *all* sensations, internal and external: they represent modes of extension, but in a qualitative manner contributing to the sensation's biological function. Unlike (b), (P) applies universally.

---

[77] To More, 2/5/49, AT V.271, CSM III.362.

[78] *6th Med*., AT VII.88, CSM II.60.

[79] *6th Med.*, AT VII.89, CSM II.61.

[80] Recall from Section 2: we can allow that he is aware of his mental state without making that state an object of his cognition (at least not in the same sense as its content is); so if the qualitative character is linked with the "conscious" or transparent nature of the state, he may be aware of it without its being "what" he senses.

[81] To say that senses "report the truth" here, I think, is to say that they enable us to make generally accurate judgments with respect to how to preserve our well-being. As Descartes says similarly of the passions, "their natural function is to move the soul to consent and contribute to actions which may serve to protect the body or render it in some way more perfect" (*Passions* II.137, AT XI.430, CSM I.376). This does not require that *what* sensations represent be ecological properties, only that *what* they represent be represented in a manner reflecting ecological salience. Both (P) and the bio-functional theory accommodate the making of the accurate judgments; (P) has the added advantage that our sensory states do indeed report what is in the bodies they appear to be reporting on.

It seems to me, then, that bio-function neither grounds Cartesian representation nor constitutes its content. Rather, I think, bio-function illuminates the *qualitative* character of sensory representation. Simmons herself provides the material to make this clear. Cartesian sensations, she writes, "have the representational function of acquainting the perceiver with features of the world in a way that facilitates self-preservation;"[82] they "need to acquaint us with the corporeal world in such a way that we can make appropriate (and often quick) judgments on their basis about how to act; this means representing it perspectivally … and in a motivationally effective way."[83] The "idea here is that there is something about the way sensations represent their objects that afford them a motivational salience that clear and distinct intellectual perceptions just do not have."[84] All this may explain why, for Descartes, God endowed us with sensory states: their qualitative character moves us efficiently to reach conclusions and react in a way purely intellectual grasp of mechanical properties may not.[85] But it doesn't explain what makes sensations *count* as representational, nor determine just what their particular representational contents are. Or put differently: It may help explain why God constructed us as He did, but it does *not* provide a theory of content.[86] Once the bio-functional role of sensations is satisfied by their qualitative character, there's just no reason to treat their representational content as performing that same role.

---

[82] Simmons (1999, p. 367, n. 26).

[83] *Ibid*., p. 356.

[84] *Ibid*., p. 367, n. 21.

[85] Malebranche usefully spells out some details here (*Search* I, p. 10ff. (p. 48ff.)). But the very fact that he also denies that sensations represent indicates that it is not by virtue of their representing that sensations fulfill their bio-functional role.

[86] Simmons, in places, seems to admit this point. She observes that God pairs pineal gland motions with sensations in a way conducive to self-preservation, where the sensation chosen is "one that (a) phenomenally presents to the perceiver (what is typically) the original distal cause of the pineal motion and (b) phenomenally presents that cause in a way that permits the perceiver to engage in appropriate interactions with it." (p. 357) This observation indicates, as noted above, that the *content* of the sensation precedes its functional role, as does the qualitative character of that content. Further Simmons admits as much by next noting that God's choices would "make little sense if sensation-types were intrinsically vacuous." (p. 358) But rather than see this as a problem for a bio-functional account of representation, Simmons here distinguishes "seeming to represent" from "actually representing," where the latter requires being "locked into a causal and ecological system." (p. 358) But this move—a version of Wilson's "presentational"-"referential" distinction—is not very plausible, for many of the same reasons. In particular, what's central is that Cartesian mental states represent whether or not they succeed in denoting some actually existing object; Simmons cannot argue that bio-functional role is necessary for representation by *stipulating* that a state has content only when it is actually "locked" into a causal system.

5.3 Intrinsic, Primitive Sensory Representation

So neither a causal, nor bio-functional, theory accounts for Cartesian sensory representation. Since these are likely the best externalist theories of representation, we should expect that, for Descartes, that, and what, a sensory state represents depends only on its intrinsic properties. Or put differently: If we hope to get a handle on Cartesian representation, we must focus on presentational content, which despite its loose definition at least restricts itself to the intrinsic properties of the sensory state, viz. such properties as being modes of mind, of being conscious states, of having a qualitative character, etc. Admittedly, none goes very far towards explaining sensory representation, nor does Descartes attempt any such explanation. But this is just as we might expect: Cartesian sensory representation is not merely intrinsic, I think, but also a primitive.

(1) First, consider intellectual ideas, his representational paradigm. Neither externalist theory had any purchase here. And while intellectual representation is not controversial, Descartes gives no explanatory account of it, of just how objects come to "be" in the intellect.[87] Further, we saw that objective being was an "intrinsic denomination" of the object, and that "…the objective mode of being belongs to ideas by their very nature …."[88] These strongly suggest that Descartes takes intellectual representation to be intrinsic and primitive. If he accepts it here, at least the burden of proof must be on those who think he demands something more for sensory representation.

(2) Cartesian dualism treats thought and extension as "simple": there is nothing more basic from or out of which they are constituted.[89] If the essence of the mental is to represent—as suggested earlier—then we should expect representation to be a Cartesian primitive, and in a way that sharply distinguishes it from anything physical.

Indeed the crucial error with externalist theories is their failure to respect this point, by neglecting a distinction we might make between a "sign" and a genuine "representation." A "sign" is a physical entity which does or can cause a mind to enter a representational state. Descartes writes:

---

[87] Cf. Gueroult: That ideas represent is a "first declaration (*une constatation première*)" of the Cartesian system; how ideas represent, by possessing objective reality, and how the mind processes these are "questions that Descartes has not seen fit to ask or to resolve, because, according to him, they exceed the limits of our capabilities … [Representation] is a first given that is revealed to us by natural light, before which every investigation stops." (Gueroult 1952/1984, p. 91)

[88] *3rd Med.*, AT VII.42, CSM II.29.

[89] *Rules* 12, AT X.419, CSM I.44. And while he also suggests that the mind–body union, the locus of sensation, is a simple or primitive notion (To Elizabeth, 6/28/43, AT III.691, CSM III.226), it's clear that the categories of "purely intellectual or purely material," both knowable via the intellect, are his most basic ontological ones. (See O'Neil 1974, Ch. 1, for discussion.)

> Words … bear no resemblance to the things they signify, and yet they make us think of these things … Now if words, which signify nothing except by human convention, suffice to make us think of things to which they bear no resemblance, then why could nature not also have established some sign [*signe*] which would make us have the sensation of light, even if the sign contained nothing in itself which is similar to this sensation? Is it not thus that nature has established laughter and tears, to make us read joy and sadness on the faces of men?[90]

In general, Descartes compares nerve or brain states to signs as well.[91]

Note, however, that accounts are on offer of that in virtue of which all these signs designate: language by convention, laughter/tears by their causal relations to underlying passions, light, and brain states, by the ability to cause sensations etc. What all these share, what *makes them* signs, is their ability to cause the mind to enter representational states: words make us think of their denotations, laughter/tears make us "read joy and sadness," etc. But no account is on offer of that in virtue of which the *mental state* might be "of" the word's denotation, of the joy/sadness, etc. Nor *could* these mental states themselves be "signs" in the same sense: signs (can) cause minds to enter states which read or interpret those signs, but the interpreting states themselves are clearly not open to the same analysis on pain of a vicious regress.[92]

A "sign," then, represents derivatively, by virtue of its ability to cause a relevant mental state and by virtue of various external relations it holds to that which it designates; a "representation" represents primitively and intrinsically. Various physical entities function perfectly well as Cartesian signs, but only mental states are representations. The failure of externalist theories of sensory representation—and so of the notion of referential content—is thus their attempt to analyze representations as if they were signs.

(3) Descartes' detailed account of the etiology of sensation might suggest a theory of representation. Yet at all the crucial moments—where the physical gets connected to the mental—he promptly refers to that which is "ordained by nature."[93] Physical states lead to mental states because God made it so. Nothing here implies those relations in any way ground mental representation, as externalist theories require.

(4) Malebranche famously critiqued Descartes' claim that we have a clear idea of the mind, noting that we lack any a priori grasp of what modifications the

---

[90] *World* 1, AT XI.4, CSM I.81.

[91] *World* I, AT XI.3–6, CSM I.81–82; *Optics* IV, AT VI.112–114, CSM I.165–166; *Principles* IV.197, AT VIIIA.320–322, CSM I.284.

[92] Descartes was clearly sensitive to the regress issue: *6th Replies*, AT VII.422, CSM II.285; *7th Replies*, AT VII.559, CSM II.382; *Optics* VI, AT VI.130, CSM I.167. Cf. also note 16.

[93] *Optics*: AT XI.130, CSM I.167; AT XI.130, CSM I.169; AT XI.137, CSM I.170; *Principles* IV.197, AT VIIIA.320, CSM I.284; *Passions*: I.36, AT XI.357, CSM I.342; I.44, AT XI.361, CSM I.344; I.50, AT XI.368, CSM I.348; II.137, AT XI.430, CSM I.376.

mind is capable of, including how it represents.[94] Descartes' own, earlier explication of his claim,[95] in fact doesn't dispute that point.

(5) Finally, that representation is a primitive was commonly accepted in the period.[96] For just two examples, consider two "Cartesians" who could hardly disagree more deeply in general, yet who agreed on this. Malebranche held that God intrinsically and humans derivatively manifest representational content, but not only explains neither but explicitly denies that he should be expected to.[97] Similarly, Arnauld, insisting that humans *can* inherently manifest content, offered no account of how that is possible, and specifically denied we can model that process on anything more familiar.[98] If Descartes had a theory of representation, it's hard to believe that two of his closest readers would have missed it or failed to invoke it in their decades-long disputes about his work.

6 Recap of the Argument

(P) proposed that, despite transparency, Cartesian sensation, like intellection, is direct: Just as an object (or property) can be "in the intellect," so that non-reflective awareness of one's thinking may be identified with direct awareness of the object, so too an object can be "in the senses," so that non-reflective awareness of one's sensing may be identified with direct awareness of the object. For (P) to work, we saw, Cartesian sensations must represent intrinsically. We then critiqued the best externalist theories of representation, and argued that Cartesian sensory representation is indeed intrinsic, as well as primitive. So far, so good, for (P).

7 Objection to (P): The Resemblance Problem

Perhaps the strongest consideration against (P) and for the "veil" comes from Descartes' insistence that sensations fail to resemble that to which they may be externally related.[99] It's hard to suggest that Cartesian sensations intrinsically represent their "objects," after all, when he relentlessly stresses their differences.

---

[94] *Elucidations* XI, p. 636. Cf. Schmaltz (1996), Lolordo (2005).

[95] *2nd Med*., *3rd Replies*.

[96] Nadler (1992a): "[The] way in which an idea presents or displays a [representational] content … is basic and inexplicable …. This is true for Malebranche, and it is also true for Descartes, Arnauld, and Régis." (pp. 50–51)

[97] *Search* III.ii.5, p. 229; *Réponse*, p. 288. See Pessin (2004) for discussion.

[98] *True/False*, pp. 66–67, *VFI* 38, p. 199.

[99] *World*, AT XI.3–4, CSM I.81; *3rd Med.*, AT VII.37, CSM II.26; *6th Med.*, AT VII.81, CSM II.56; *Principles* I.70, AT VIIIA.34, CSM I.218.

Similarly, if we're aware of our sensings, and these (or their apparent objects) do not even resemble that to which they may be externally correlated, then, just as argument (A) concludes, it would seem that *what* we're aware of cannot be identified with anything external.

(1) But we must be careful here, as the notion of "resemblance" in play is not apparent. No "idea" or mental state could ever literally resemble a physical state, obviously, since immaterial and material substances can share no non-generic properties. If Descartes ever speaks of any ideas "resembling" bodies,[100] or is to speak substantively of denying resemblance, he must have something else in mind.

(2) Consider Cook's (1987) suggestion that ideas resemble objects when they contain objectively what their objects contain formally.[101] Recall from Section 4 that there were two options with respect to whether sensations contain objective being: On (i) they do despite being poor epistemic resources for determining the true nature of their objects, and on (ii) they do not, because of their so being. On (ii), sensation-y would fail to resemble its object (as desired) because it fails to contain objectively the motion-m formally modifying the sun. Nevertheless, (ii) can grant (as we've seen) that sensations do intrinsically represent their objects, which is all (P) needs. Non-resemblance would thus provide no basis either for denying intrinsic sensory representation or for affirming a veil.

On (i), to the contrary, Cook's account seems inconsistent with (P). On (i), sensation-y, representing motion-m, contains motion-m objectively; conjoined with Cook's account, y would count as resembling its object. Since Descartes rejects that resemblance, we must deny that y represents m and so deny (P). But in fact the problem here is with Cook's account, for it fails to distinguish between the content of a state and its qualitative character. The preceding argument assumed that Descartes' rejection of resemblance concerned the former only and not the latter. But suppose it actually concerns both.[102] Sensation-y has as its content motion-m and has its yellow qualitative character. Cartesian ideas resemble their objects, I propose, when (i) they contain objectively what their objects contain formally *and* (ii) when their qualitative character does not in some way obscure this content from the cognizing mind. The problem with y, of course, is that it fails the second condition. But this result is consistent with its fulfilling the first condition, and having, as its object, m—which is all (P) needs.

On this reading, note, intellectual ideas may resemble their objects while sensations of secondary qualities, at least, will not. But all that ultimately means is that sensations simply fail to reveal their objects' "true" (i.e., mechanical) natures.

---

[100] *Principles* II.1, AT VIIIA.40, CSM I.223; *3rd Med.*, AT VII.39, CSM II.27.

[101] Possible supporting texts: *2nd Replies*, AT VII.161, CSM II.114; *3rd Med.*, AT VII.41–42, CSM II.29.

[102] That Descartes *would* construe (non)resemblance in terms involving the qualitative character of sensations is very plausible. His main concern is that how the world appears to us via the senses might not be how it really is, and the "appearance" element is precisely what is reflected in sensations' qualitative character.

That's perfectly consistent, again, with saying that *what* Fred senses is m, hence that y intrinsically represents m, hence that y just is m, sensed, as (P) requires.

(3) Similar considerations apply to a related suggestion, that an idea resembles its object insofar as its object conforms to the idea, or the idea is "true" to its object.[103] I assume that, for Descartes, an idea is true to its object if and only if the right sort of correspondence obtains between them.[104] Since Cartesian sensory differences "correspond" to material differences despite not resembling them,[105] I also assume that the "right sort" of correspondence involves more than just some kind of externalist mapping. Now, again, an "idea" offers us two possibilities: On (P), the object of sensation-y, viz. motion-m, certainly "corresponds" to m; on this reading, y would be true to its object, hence resemble it, contra Descartes. So suppose that what's in play again is instead y's qualitative character, its yellowness. This may fail to "correspond" to m in the sense desired, thus ensuring that y is not "true to its object," hence fails to resemble it, but all in a way consistent with (P). Again, then, non-resemblance would be compatible with (P).

But we might go still further. Strictly speaking, the object of sensation-y, on (P), just is motion-m; y just *is* m, sensed. The problem is that y on its own fails to allow Fred to determine the true nature of y's object. Suppose, then, that the "right sort" of correspondence were partly governed by epistemic needs. If so, then—providing a small gloss on the account of resemblance in (2)—an idea might be "true" to its object insofar as, on the basis of the idea alone, its agent may determine the nature of its object; if so, an idea would *resemble* its object if and only if the true nature of its object were epistemically accessible to the agent on the basis of the idea alone. (P) would say just this: sensation-y fails to resemble motion-m insofar as we cannot determine visually that its object is m. But that is perfectly consistent with m's actually being y's object. An intellectual idea, meanwhile, might resemble its object precisely because, with attention and reflection to it, we *can* determine quite precisely the true nature of its object.

There is much to support this suggestion:

(i) First, Descartes sometimes implies not that there are no colours (etc.) in the physical world but rather that colours are to be identified with physical properties.[106] If so, then there *are* colours in the world, only our sensations fail to reveal their true nature; or put differently: our sensations *are* "of" (physical) colours, although we cannot determine on their basis that this is so.

---

[103] This proposal is also in Cook (1987, p. 187), who treats it as equivalent with the preceding proposal. Cf. To Mersenne, 10/16/39, AT II.597, CSM III.139; *3rd Med.*, AT VII.37, CSM II.26; *World* 1, AT XI.5, CSM I.82.

[104] Cf. Frankfurt (1978): "… [W]henever Descartes gives an explicit account of truth he explains it unequivocally as correspondence with reality." (p. 37) Although truth/falsity might, strictly, be restricted to judgments, they also have a ready application to ideas.

[105] *6th Med.*, AT VII.81, CSM II.56.

[106] *Principles* I.70, AT VIIIA.34–35, CSM I.218; IV.198, AT VIIIA.322–323, CSM I.285; *World* 2, AT XI.9, CSM I.84; *Rules* 12, AT X.413, CSM I.40–41.

(ii) More generally, Descartes' main complaint about sensations is that they are poor epistemic resources, in particular compared to intellectual ideas. He regularly describes them in epistemic terms such as "obscure" and "confused,"[107] saying that "there is no way of understanding what sort of things they are. If someone says he sees colour in a body…this amounts to saying that he sees or feels something there of which he is wholly ignorant."[108] He also notes that we know what colours are in "quite a different way" from the way in which we know what size, size, and motion are—qualities he says are "clearly perceived"[109] etc. More specifically, as Wilson observes, there is a "close association, in Descartes' writings, between saying sensations are 'confused' or 'obscure', and insisting that they fail to 'resemble' external things."[110] She writes: "To say that sensations are 'confused' or 'obscure' is to say that they fail to provide a distinct *understanding* of real qualities of bodies …. A 'non-resembling' idea … should be construed as one that fails to yield intelligibility."[111] In my terms, the non-resembling sensory idea fails to yield intelligibility precisely because we cannot determine, on its basis, the true nature of the bodily properties which are its object.

(iii) Most generally, Wilson also observes that Descartes illuminates his denial that mental states need resemble the objects they represent by "considering alleged non-resemblance between cause and effect *within* the realm of the mental: the thought of a tree is not at all like the experience of hearing the word 'tree', which brings it to mind."[112] She next quotes this text:

> Most philosophers maintain that sound is nothing but a certain vibration of air which strikes our ears. Thus, if the sense of hearing transmitted to our mind the true image of its object then, instead of making us conceive the sound, it would have to make us conceive the motion of the parts of the air which is then vibrating against our ears.[113]

This leads Wilson to propose that "Descartes' denial of 'resemblance' between his sensory ideas and the qualities of bodies … relies to some extent on comparing mental awarenesses. Against the experience of the sensation of sound, we place the (mental) 'image' of motions of air particles. The two, phenomenally, do not 'resemble' each other."[114]

So, we determine whether a given mental state is "true" to its object by applying another mental state or faculty to it; determination of the true nature of the object of sensation is itself a matter of intellectual reflection. But now, how can the intellect perform this evaluation if not by, in essence, comparing the

---

[107] *Principles* I.197, AT VIIIA.320, CSM I.284; *3rd Med*., AT VII.43, CSM II.30.

[108] *Principles* I.68, AT VIIIA.33, CSM I.217; I.70, AT VIIIA.34, CSM I.218.

[109] *Principles* I.69, AT VIIIA.33–34, CSM I.217.

[110] Wilson (1994, p. 22). See *Principles* I.70, AT VIIIA.34–35.

[111] Wilson (1994, p. 22).

[112] *Ibid*., p. 20.

[113] *World* 1; AT XI.5, CSM I.82. Cf. letter to Regius, September 1642, AT III.493, CSM III.206.

[114] Wilson (1994, p. 20).

sensation with its own "image" (or "conception")? What I'm suggesting then, is that sensations' not resembling their objects amounts, ultimately, to their being epistemically opaque *to the intellect* in a way that intellectual ideas are not. But this is an epistemic difference: intellectual ideas can be clear and distinct, are subject in a particular way to rational evaluation and justification, are susceptible to the truth rule, etc., while sensory ideas are not. But then—again—this difference does not imply that sensory states fail to represent, fail to possess objective being, etc. It merely implies that sensory states cannot inform *the intellect*, in the right way, *that* they do represent, possess objective being, etc. Sensation-y can just be motion-m itself in Fred's senses, therefore, just as a certain conception of that motion just is m itself in Fred's intellect, even if y is epistemically opaque to the intellect. It is thus no accident that Descartes notes that "…we cannot find any *intelligible* resemblance [*nec ullam similitudinem intelligere possimus*] between the colour we suppose to be in objects and that which we experience in our sensation[:]"[115] the lack of resemblance is constituted by the intellect's inability to grasp the actual object of the sensation, and not by the sensation's lacking an object. That sensory ideas don't resemble their objects, so construed, again is no obstacle for (P).

(iv) We're now in a position to patch up Wilson's "presentational" content, which she over-loosely defined as "what the mind takes itself to be aware of." This definition allowed or overlooked a possible gap between what the mind takes itself to be aware of and what it is *in fact* aware of. If in light of that gap we still retain Wilson's definition then "presentational" content will be (say) the content of an intellectual state reflecting on a sensory state; given the "lack of intelligible resemblance" etc. the intellect may well judge that the sensory state has *no* content, and is "merely" a state of mind, or else judge that its content is so vague or obscure that one might say that "the reality which [sensations] represent is so extremely slight that I cannot even distinguish it from a non-thing,"[116] etc. On this view we can understand many scholars' temptation to read Cartesian sensations as being non-representational, and so to support the veil. The fundamental problem with such a reading, now, is that its conclusion about sensation is based on the intellect's judgement about sensory states, but *not on the nature of the sensory states themselves*. Now given everything we've said, it's possible that the intellect is here making a mistaken judgement about sensations; or less strongly, that sensory states are simply opaque to the intellect in this way. If so, then "what the mind takes itself to be aware of" during sensation might not, after all, be what the mind is *in fact* aware of. All those texts suggesting sensations are non-representational are thereby disarmed.

I propose a terminological change: let us apply the term "presentational content" simply to "what the mind *is* aware of." So defined, presentational content just is sensory content, which, as I've argued, is intrinsic, and primitive, and

[115] *Principles* I.70, AT VIIA.34, CSM I.218, my emphasis.

[116] *3rd Med*., AT VII.44. CSM II.30.

opaque to the intellect. So construed, there is nothing left in presentational content to support, or even to imply, the veil of perception.

(4) Finally, the preceding indicates that the non-resemblance issue is in fact a red herring with respect to (P). All (P) needs is that sensation-y intrinsically represents motion-m. Resemblance would be relevant only if Descartes also held that resemblance were necessary for representation. But as often as Descartes denied that sensations resemble bodies, he also denied that resemblance is necessary for one thing to be a "sign" of another, and often in contexts in which he was comparing representations *to* signs;[117] and while his many examples of signs and representations not resembling their objects perhaps suggest the differences between sensations and their objects, they equally well suggest that sensations are perfectly legitimate representations of their objects *despite* their non-resemblance.[118] This is particularly true given (from above) that Cartesian representation is a primitive: since it's not grounded in any particular features of the sensation, it need not require the sensation's resemblance to its object.

In short, then: The non-resemblance issue provides no obstacle for (P). To the contrary, given the above, it actually coheres quite nicely with (P).

8 Descartes' Quining of Qualia

> Even if I do not refer my ideas to anything outside myself, there is still subject-matter for error, since I can make a mistake with regard to the actual nature of the ideas. For example, I may consider the idea of colour, and say that it is a thing or quality; or rather I may say that the colour itself, which is represented by [*per*] this idea, is something of the kind. For example, I may say whiteness is a quality; and even if I do not refer this idea to anything outside myself—even if I do not say or suppose that there is any white thing [*ac dicam vel supponam nullum esse album*]—I may still make a mistake in the abstract, with regard to whiteness itself and its nature or the idea I have of it ….[119]

---

[117] *World* 1, AT XI.4, CSM I.81; *Principles* IV.197, AT VIIIA.320–321, CSM I.284; *Optics* 4, AT VI.112–114, CSM I.165–166.

[118] Simmons (1999) makes something like this point. She also points out that Descartes is concerned to reject specifically the scholastic Aristotelian account of sensory representation, which relied on "intentional species" assumed to represent their objects by virtue of resemblance—but "this claim is not a denial that sensations represent anything in corporeal reality." (p. 351) Ironically, Descartes probably misunderstood the scholastic account here: As Pasnau (1997, Ch. 3) argues, Aquinas' (and others') stress on the "likeness" between knower and known, and so between objects and the "species" they transmit, was meant merely to deny that cognitive representation was conventional. Further, the "likeness" was satisfied by virtue of object, species, and the cognizing mind coming to realize the same "form," even though its realization in the species and mind didn't turn those entities into instances of the originating objects—in other words, the same form could be realized in these entities without their "resembling" each other in any literal way other than the sharing of form. So construed there is nothing here that is essentially different from Descartes' own account of cognition.

[119] *Burman*, AT V.152, CSM III.337.

On the Cartesian ontology, everything that exists is either substance or mode, or "thing" or "quality."[120] This interesting passage from *Burman* possibly suggests that colour is neither.[121] It's easy to read the passage merely as a standard Cartesian denial that colour is a quality of bodies, or mode of extension. But note: he specifically denies that colour is a quality even when he is *not* referring to anything "outside" himself. This may suggest that he denies that colour is a quality *simpliciter*, a thesis with substantial exegetical implications. I won't rest anything on the interpretation of this unique and obscure passage, but it does inspire the following considerations.

If colour—phenomenal colour, colour as experienced by us during sensation—is anything, it would be a mode. Qua mode, however, it exists formally in neither matter (not being a mode of *res extensa*) nor mind (since minds are not themselves coloured). But since matter and mind exhaust the ontological possibilities, colour cannot exist formally *simpliciter*. The obvious proposal is that colour exists "objectively," "in" a Cartesian mind, despite the mind itself not *being* coloured.[122] Many scholars probably incline towards such a view, at least implicitly; in so doing they probably understand colours (and other secondary qualities) as the "objects" of sensation, as the (Wilsonian) "presentational" contents of sensation, as "what" the mind is sensing, etc., and so are easily led to the veil of perception. But my account of Cartesian sensation now moves away from this view. To state it simply: we ought not, in fact, to treat phenomenal colour as the object of Cartesian sensation.

For Descartes, it would seem, only what can, at least *possibly*, exist formally, is capable of existing objectively.[123] MacKenzie refers to this as an "objective reality principle," which she states this way: "Only those properties which can inhere formally in some substance can inhere objectively in a thinking substance."[124] Our discussion in Section 2 suggested as much: since, during representation, the objectively existing thing is the same "thing itself" as the formally existing thing, whatever exists objectively must either exist, or at least be capable of existing, formally, or be constructible from elements which are.[125] Given this principle, and the impossibility of phenomenal colour's existing formally just sketched, then we cannot treat it as existing objectively, thus as the object of sensation.

---

[120] *Principles* I.48, AT VIIIA.22–23, CSM I.208. Descartes' reference to eternal truths here is irrelevant here.

[121] There are of course also passages where Descartes refers to secondary qualities as "qualities." (To Chanut, 2/26/1649, AT V.291, CSM III.369) But in such passages Descartes may merely be speaking loosely.

[122] Cf. Hatfield (1990, p. 53).

[123] Cf. *2nd Replies*, AT VII.166, CSM II.117.

[124] MacKenzie (1989, p. 182). Cf. Normore (1986): "Thus if an idea has objective reality, and is thus of a thing, that thing possibly exists." (p. 238)

[125] Chimerae and the like could of course be constructed from simpler, "possible" elements.

An objection may arise: The "objective reality principle" is particularly evident with respect to clear and distinct perception, which always involves "true" ideas, i.e., ideas of possible things,[126] which suggests perhaps that it needn't hold with respect to opaque perception, of which sensation is the paradigm. In reply, though, note that Descartes writes that

> …pain and colour and so on are clearly and distinctly perceived when they are regarded merely as sensations or thoughts. But when they are judged to be real things existing outside our mind, there is no way of understanding what sort of things they are.[127]

Qua modes of mind colour sensations *are* clearly and distinctly perceived, and indeed, as required by the principle, the mental modes which are colour sensations are formally possible. Qua representations, however, colour sensations indeed are opaque; but on its own this fact simply means that we just cannot judge *either way*, on the basis of our sensations, whether colour is formally possible, as we saw in Section 7. Consequently there is nothing here to upset the "objective reality principle." And while the opacity of sensations means we also cannot judge on their basis alone that colour is formally *im*possible, we do have more than just our sensations to work with: Descartes stresses repeatedly that we clearly and distinctly grasp the essence of matter qua extension, and this clear and distinct grasp rules out treating phenomenal colour as a possible formal mode of matter. Since there is no inclination here to treat it as a possible formal mode of mind, we again have our conclusion that phenomenal colour does not exist objectively in the sensing mind, and so is not the object of sensation.[128]

If so, then phenomenal colour exists neither formally nor objectively. But then *it is nothing at all*, as the *Burman* text above may be suggesting. Assuming that similar considerations apply to all the other standard secondary qualities, then Cartesian qualia, as it were, are *nothing at all*. If they are nothing at all, then they certainly are not the objects of cognition. And if not, they can provide no veil.

But wait—it just is not easy to deny the reality of phenomenal colour, or our awareness thereof. Fred sees a particular surface texture "as" yellow; just what is it he seeing this texture "as"? Here I can only give the same answer as above: it's

---

[126] *6th Med*., AT VII.78, CSM II.54.

[127] *Principles* I.68, AT VIIIA.33, CSM I.217.

[128] Alternative ways of framing this whole argument are available. We might ask whether, for example, phenomenal colour is to be identified with sensory ideas "taken formally (or materially)" or "taken objectively." (See note 26.) The former is a non-starter; but to "take an idea objectively" is to consider it in relation to the thing it represents. But since neither mind nor matter are actually, formally, phenomenally coloured, and these exhaust the possible *representanda*, we cannot identify phenomenal colour with a sensory idea taken objectively. One might suggest, as Normore (1986) does, that sensations are the sorts of things for which the formal/objective distinction doesn't arise: in effect, sensory ideas represent, or are "of," themselves, so that the idea of warmth (for example) just is warmth itself. The problem with this suggestion is this: If a sensory idea represents itself, and (as we just noted) we clearly and distinctly grasp that it is formally a mental mode, then qua representation (of itself) it ought to be clear and distinct. But it is precisely the sensory idea qua representation which Descartes holds to be opaque.

a primitive Cartesian fact that this is "what it's like" to be in that mental mode which is the sensing of motion-m. But that this mode has this qualitative character is not any further fact or property really distinct from the mode itself, any more than its being a (non-reflectively) conscious state is (or rather, Fred's being conscious of it is): All we have are Cartesian mental states which are just the sorts of things which are conscious, have contents, and qualitative characters. No further information is available about them because there is none. Why does sensing m give rise to this yellowishness? That is just what it is to consciously sense m. One must, above all, avoid reifying these features, as if they were really distinct from the mental states themselves. "What" we sense are certain surface textures. *Phenomenal colours just are how those surface textures look to us*, or what it is like to visually sense surface textures.

Or to be more precise, and to link this to a point in Section 2: There we saw that mental transparency, our being consciously aware of a cognitive state, did not require treating that cognitive state as an object of cognition. What I'm suggesting now is that the qualitative character of a sensory state just *is* its being a conscious state, or its *way* of being a conscious state. That we are aware of these characters, then, would not require that they constitute an object of sensation. And that's what it might mean to consider colours, and other sensory qualities, to be "nothing at all." If so, then we should resist the temptation to say that phenomenal colour itself is "what" we sense. There just is no room for it in the available ontology. If we are to sense colour at all, it must be "physical colour," i.e., the relevant surface textures, such as motion-m, just as (P) proposes (cf. 8.3.1 above). Of course we don't sense physical colour "as it is"—as it is conceived by the intellect—but it is still physical colour we sense. There just isn't anything else available *to* sense. Again, there is no veil.

A final, highly speculative point. I've called this section "Descartes' Quining of Qualia." This title is inspired by Dennett (1988), which borrows the definition of "to quine" from *The Philosophical Lexicon*, a satirical dictionary of eponyms: "To deny resolutely the existence or importance of something real or significant."[129] One might object that in this section I've not exactly quined Cartesian qualia, since my denial of the reality of phenomenal colour comes at the cost of accepting the "qualitative character" of sensory experience, with which contemporary philosophers, in fact, often identify qualia. True enough. But I'll close this section by suggesting that Descartes might have had a further project brewing, vaguely, in some remote corner of his mind: that of eliminating qualitative character from his ontology altogether, or at least reducing it to other less controversial things (for him), such as intentional states endowed with a variety of causal properties.

Consider this text from Hobbes' *Third Objections*:

> Even if we grant that fear is a thought, it can only … be the thought of the thing we are afraid of. For what is fear of a charging lion if not the idea of a charging lion plus the

[129] Dennett (1987).

> effect which this idea produces in the heart, which in turn induces in the frightened man that animal motion which we call 'flight'?[130]

Descartes' immediate curt reply—"It is self-evident that seeing a lion and at the same time being afraid of it is different from simply seeing it"—does not actually reject Hobbes' proposal, which is aimed merely to deny that fear (and other states) are peculiarly mental in nature. This suggests that perhaps the intentional state of fearing that p might be analyzed, ultimately, not as involving the representation of p plus some irreducible "qualitative character" of fear, but rather as the representation of p plus certain kinds of consequent behavioral dispositions.[131] If such a project were successful, and could be generalized to all sensory states, then one would no longer have to treat qualitative character as a primitive. Descartes obviously did not develop any such theory, but, maybe, just maybe, the seeds are there.[132]

Conclusion: The Unity of the Cartesian Mind

> … [T]he power through which we know things … is one single power …. It is one and the same power: when applying itself along with imagination to the 'common' sense, it is said to see, touch, etc.; when addressing itself to the imagination alone, in so far as the latter is invested with various figures, it is said to remember; when applying itself to the imagination in order to form new figures, it is said to imagine or conceive; and lastly, when it acts on its own, it is said to understand … According to its different functions, then, the same power is called either pure intellect, or imagination, or memory, or sense-perception.[133]

(P)'s final advantage is that it coheres well with the unity of the Cartesian mind, expressed above.

For despite Descartes' distinguishing various aspects of the mind, he never loses sight of the mind's ultimate unity.[134] This yields the following picture: The mind by its nature differs from matter in manifesting states which have, primitively, a representative nature and a conscious, (possibly reducible) qualitative character. But even this twofold nature is a unity: the qualitative character of the state just *is* the way the *representandum* is experienced in cognition. And while every mental state has both aspects, the precise relationship between them can vary. For some

---

[130] *3rd Objections*, AT VII.182, CSM II.128.

[131] We saw in Section 5 that Descartes was very aware of the close links between certain internal sensations and their causal consequences, despite not using them in the *analysis* of representational content.

[132] For development of such a theory in contemporary philosophy of mind see Dennett (1988, 1991, 2005), as well as discussions of the "Representational Theory of Qualia/Consciousness." (Lycan 1998, Tye 2000, Lycan 2004, Jackson 2004) Proper evaluation of whether Descartes might incline in these directions would require close study of *Passions* in particular.

[133] *Rules* 12, AT X.415–416, CSM I.42.

[134] *6th Med.*, AT VII.85–86, CSM II.59.

states, the qualitative character reaches some maximum degree of intensity obscuring epistemic access to the objective being contained, while for others it reaches a minimum degree allowing clear and distinct access. States with a higher qualitative intensity may thus wrongly be taken not to represent; states with a lesser may (perhaps) wrongly be taken to lack a qualitative nature.[135] Obviously "sensory" states fall under the former, while "intellectual" fall under the latter.[136]

To be sure, there are important differences between these as well. They are generated by different processes. Sensory states arise, as Descartes notes, when the mind applies itself to the "common sense," while intellectual states arise when the mind "acts on its own." Thus sensory states are oriented towards the body, or the mind–body union, in a way that intellectual states aren't.[137] Thus sensory states serve different goals from intellectual states: fostering the mind–body union versus aiming at "truth," respectively. The former generally requires (as we saw) quick and motivationally effective responses to the world; the latter requires reflection and reasoning, hence time. Thus we have the Cartesian explanation, discussed above, for why God might have endowed us with states varying in their qualitative, hence epistemic, hence pragmatic character.

But while the manner of representation varies according to etiology and function, with varying epistemic results, all mental states are representational, and are so in just the same sense. A mental state primitively, intrinsically, contains its object; the object exists objectively in that state. When sensory processes result in Fred's sensing motion-m, then motion-m is "in his senses" in that relatively obscure way just described; when intellectual processes result in his thinking of m, then m is "in his intellect" in that clear way. Since mental transparency also dictates our primitive awareness of our mental states, in being aware of his sensory state or in being aware of his intellectual state, Fred in fact is aware of motion-m *itself*, as it is sensed, or as it is thought. Motion-m is the object of his cognition in both cases. Rather than undermine it, then, Descartes' doctrine of transparency supports the directness of both Cartesian sensation and intellection, and they all cohere beautifully with the unity of the Cartesian mind.[138]

---

[135] "Wrongly," because there is "something that it's like to think of x," even though that something is not as vivid or intense as what it's like (say) to see yellow.

[136] This framework is applicable to the imagination and the passions as well.

[137] *Principles* I.48, AT VIIIA.23, CSM I. 209. Simmons (1999) explores this thoroughly.

138 Thanks to many people for helpful comments on this paper. Versions or excerpts of this paper have been presented at the New England Colloquium for Early Modern Philosophy, the University of Western Ontario, Wesleyan University, Connecticut College, and Queen's University. Thanks go to those audiences in general, and to individuals in particular including Alison Simmons, Justin Broackes, Jeff McDonough, Tom Lennon, Steve Horst, Sanford Shieh, Melvin Woody, Kristin Pfefferkorn, Larry Vogel, Derek Turner, Simon Feldman, Jon Miller, Charlie Jarrett, Marleen Rozemond, Don Garrett, and Brian McMaster. Several people also provided very helpful written comments, including Kurt Smith, Larry Nolan, Raffaella De Rosa, and Paul Hoffman. Finally I'd like to thank Queen's University and Connecticut College for supporting my travel to present this paper, and especially to thank the students in my Descartes seminar at Conn for helping me work through a number of the ideas therein.

Abbreviations

AT = Oeuvres de Descartes, eds. Charles Adam and Paul Tannery (Paris: 1897–1910 and 1964–1978; Paris: Librairie Philosophique J. Vrin, 1996).
Burman = Conversation with Burman, in *CSM III.*
CSM = The Philosophical Writings of Descartes, vol. I, II, transls. J. Cottingham, R. Stoothoff, and D. Murdoch, and *vol. III*, transls. J. Cottingham, R. Stoothoff, D. Murdoch, and A. Kenny (Cambridge, UK: Cambridge University Press, 1984, 1985, 1991).
Discourse = Discourse on the Method, in *CSM I.*
Elucidations = Elucidations of The Search After Truth, transl./ed., Thomas Lennon, in *Search*.
Med. = Meditations on First Philosophy, in *CSM II.*
Meteorology = Les Meteores, in *AT VI*.
OA = Oeuvres de Messire Antoine Arnauld (Paris: Sigismond D'Arnay et Compagnie, 1780).
Objections = Objections and Replies, in *CSM II.*
OC = Oeuvres Complètes de Malebranche, ed. André. Robinet, 20 volumes (Paris: J. Vrin, 1958–1967).
Optics = Optics, in *CSM I.*
Passions = The Passions of the Soul, in *CSM I.*
Principles = Principles of Philosophy, in *CSM I.*
Replies = Objections and Replies, in *CSM II.*
Réponse = Réponse du P.Malebranche à M. Regis, *OC 17-1*.
Rules = Rules for the Direction of the Mind, in *CSM I.*
Search = The Search after Truth, transls./eds. Thomas Lennon and Paul Olscamp (Cambridge, UK: Cambridge University Press, 1997).
True/False = On True and False Ideas, transl. Stephen Gaukroger (Manchester, UK: Manchester University Press, 1990).
VFI = *Des Vraies & des Fausses Idées*, in *OA* 38.
World = The World, in *CSM I.*

Wonder Among Cartesians and Natural Magicians

Brian Jonathan Garrett

"If philosophy is the removal of wonder, as Aristotle said, its proper enterprise is explanation." —Bas van Fraasen, 2004, p. 132.

"God made three marvels: something out of nothing, free will, and God in man." —Descartes, AT X.218.

Introduction

In the late Renaissance, the power and presence of wonder couldn't be denied; the hundred years prior to Descartes' youth might well have been labeled an Age of Wonders.[1] Europe had been inundated with wonders: the New World was plentiful with peoples and civilizations, exotic animals, and wonderful new simples, none of which was mentioned in the Bible, neither by the ancients. Indeed, the rhetoric of wonder appears in a large number of books in the 16th and 17th centuries—texts often purporting to be of practical use, divulging secrets of old, and experiments and observations of the new.[2] Wonder was found in the newly recovered ancients, perhaps as much as in the encounter with the New World; the old had become novel, so that it too would uncover its secrets. And then there were the recent marvels of technology—the mechanical clock, the telescope, the microscope, the *camera obscura*, and the experiments with the loadstone. By the late17th century Europeans had become amazed b y their own works—such as the self-moving clocks and "speaking" fountains—and there were plenty of people eager to contribute to this practical knowledge and its wonderful display.[3]

But Descartes and Malebranche are not enthusiastic about wonder, despite its obvious attractions. For although wonder's power and presence couldn't be denied, it was also a distraction, and could be quite misleading, encouraging the wrong kind of science. Wonder has its place as an incentive to inquiry, and for the embodied soul is perhaps a necessary requirement for inquiry. As a passion,

[1] Daston and Park (1998, p. 171) suggest this.

[2] Eamon (1993), Daston and Park (1998), Greenblatt (1991), Shumaker (1989), Terpak (2001).

[3] Perhaps the best known of the 17th century wonder propagators was the Jesuit philosopher Athanasius Kircher (1602–1680) and the Natural Magicians, such as Giambattista della Porta (1535–1615) and Robert Fludd (1574–1637). Indeed talk of wonders and how to (re-)create them was common among so-called "Natural Magicians."

J. Miller (ed.), *Topics in Early Modern Philosophy of Mind*, Studies in the History of Philosophy of Mind 9, 39–55.

used correctly, it will contribute to the moral and intellectual health of the individual. But just as any passion might, it could mislead by being misused. Wonder is important to Descartes' picture of the intellectual, especially in that generosity (or magnanimity) is derived from wonder, and it is generosity that is the key to the good life, i.e., to the proper control of the passions. Examining the idea of wonder thus reveals how Descartes and Malebranche differ from their contemporaries regarding the role of wonder in inquiry and, I shall suggest, religious devotion.

The following discussion is designed to explain, by brief comparison of Descartes' views with his contemporaries, why Descartes and Malebranche are not overly enthusiastic regarding wonder and its role in the life of inquiry. I shall illustrate some of the main features of wonder, along with a discussion of the main uses and misuses of wonder.[4] I situate Descartes' discussion of wonder among his contemporaries, in particular, Jesuit thinkers and "Natural Magicians", to show how Descartes' and Malebranche's views of wonder can be contrasted with those who do not share, what we might call, the same intellectual "ethics" of wonder. Wonder's link to reverence and to esteem was used by many to encourage religious devotion. To wonder at God's creation was a natural response to such a great mystery. Although Malebranche indulges in the rhetoric of wonder, especially at the marvelous body–machine and how its passions are well-adjusted, arranged for human society and individual well-being, he doesn't explicitly advocate the use of wonder as an encouragement towards religious devotion. Nor does Descartes.

The paper begins with an aside—a brief digression into an anecdote of the New World and the wonders of mechanism and life. In Section 2, I turn to an account of wonder in Descartes' *Passions of the Soul* and in Malebranche's *Search After Truth*, and in Section 3 I explore how wonder fits into Descartes' moral philosophy. In Section 4 I outline how Descartes conceives of wonder's role in inquiry and science, and in Section 5 I compare Descartes' uses of wonder with those of his contemporaries, showing that wonder encourages the wrong kind of science, one overly concerned with the sensible qualities of things. Although it is dangerous to argue from what is *not* being said, Descartes omits any comment from the *Passions* regarding how wonder at nature could encourage religious devotion. This omission seems to be significant. Wonder towards the natural world cannot be sustained along-side a desire to explain nature, so it is not possible for the optimistic rationalist to advocate wonder at nature as a means to devotion. At best, this religious use of wonder can be maintained only through self-wonder: wonder is legitimate towards ourselves, insofar as we are beings with free will, but no longer legitimate (except as a momentary encouragement to inquiry) towards nature herself.

[4] Although I offer a reading of Descartes' text I consider this paper to be somewhat more broadly historical than analytic-interpretative. Close and detailed readings of *The Passions of the Soul* can be found in Brown (2006).

1 Anecdotes of Wonder, Life and Mechanism

Poor Gulliver, lost in a new and bizarre world, tied down and suffering the indignity of the Lilliputians rummaging through his clothes. And what they discover is a puzzling marvel. As the Lilliputians themselves report:

> Out of the right Fob hung a great silver chain, with a wonderful kind of Engine at the Bottom. We directed him to draw out whatever was at the End of that Chain; which appeared to be a Globe, half Silver, and half of some transparent Metal; For on the transparent Side we saw certain strange Figures circularly drawn, and thought we could touch them, until we found our Fingers stopped with that lucid Substance. He put this Engine to our Ears, which made an incessant Noise like that of a Water-Mill. And we conjecture it is either some unknown Animal, or the God that he worships: But we are more inclined to the latter Opinion, because he assured us (if we understand him right, for he expressed himself very imperfectly) that he seldom did any Thing without consulting it. He called it his Oracle, and said it pointed out the Time for every Action of his Life.[5]

Swift is a keen observer of his times and his joke trades on at least two observations. The clock had completely transformed daily life for the early modern. Fifty years earlier few people had a pocket watch, but by Swift's day they were common among the wealthy, and town clocks had become ubiquitous. The watch had indeed become a god, given how it was transforming European life. But interestingly, Swift also has the Lilliputians speculate that the watch was an animal, alive in its self-sufficient activity. This latter trope was a familiar one among the philosophers: machines can imitate life and the unwary, like the ignorant Lilliputians, can find themselves marveling and wondering at the machine, meanwhile thinking it lives. Years earlier, Robert Boyle mentions a similar case.[6]

In his *Disquisition About Final Causes*, published in 1689 (much of which was written somewhat earlier), Boyle indulges the reader with wonders and tales of far-off lands. Although advocating, against the Cartesians, a more liberal acceptance and recognition of final causes—arguing that final causes are surely required in order to guide our study of nature, whether in optics or physiology—Boyle cautions against the devaluing of proper inquiry that might result. In his last chapter, intended to establish, "(t)hat the naturalist should not suffer the Search or the Discovery of a Final Cause of Nature's Works, to make him Undervalue or Neglect the studious Indagation of their Efficient causes,"[7] Boyle compares Englishman to Chinese:

> A country fellow here in England knows something of a watch, because he is able to tell you, that 'tis an Instrument that an Artificer made to measure Time by: and that is more

[5] Swift (1726).

[6] I doubt, however, that Boyle was Swift's immediate source in this matter.

[7] Boyle (1688, p. 229).

> than every American savage would be able to tell you; and more than those Civiliz'd Chineses knew, that took the first Watch the Jesuit brought thither, for a Living creature.[8]

Descartes' follower, Robert Desgabets (1610–1678) also knew the story, writing:

> However, the same affront to reason and philosophy is committed by the Americans and the Barbarians of the Orient who not being able to understand the mechanical reasons for the movement of clocks, or the true causes of natural effects, attribute souls and intelligence to machines, and likewise to fire, lakes etc., and in doing so expose themselves to the mockery of Europeans.[9]

The story of the unwary being astounded at a mechanism and taking the mechanism as a living being could well have been a thought experiment dreamt up by one of those "New Philosophers". Descartes' treatise *The World* (1632) comes quickly to mind. In *The World* we are asked to indulge in an imaginary reconstruction of the world, a construction that leaves the appearances to be just those that we have of the bodies around us. The world will still have plants looking alive, and well, being alive; it's just that these motions are explained by Descartes' physical principles, not by an incorporeal principle or form. By the time of Descartes' late work, *The Passions of the Soul*, his mechanistic attitude to life is explicit and public. In article 6 of *the Passions* he writes:

> …(L)et us judge that the body of a living man differs from that of a dead man as much as a watch or other automaton when it is wound and contains the bodily principle of the movements for which it is constructed…(differs from) the same watch or other machine when it is broken and the principle of its movement ceases to act.[10]

But what adds an extra level of fascination to these tales is their apparent truth: Western machines were indeed mistaken for living creatures—at least momentarily—by the natives. And if these were true anecdotes, then it was an anthropological point or observation that was being brought to bear upon the epistemology of mechanism. The New World's wonders, her peoples and their reaction to the invaders, served the new philosophers well.

Boyle mentions both the "American savages" and the "civilized Chinese" to drive home his point that our ignorance over a mechanism's causes—internal mechanical causes and final cause—can lead to mistaken inferences regarding it being alive. Perhaps Boyle's version of the story of the Chinese is accurate. The Jesuits did indeed bring a clock with them to China. Mateo Ricci, who was one of the earliest to visit at the turn of the 17th century, mentions that he brought a clock as a gift, but Ricci doesn't record in his diaries any reaction by the Chinese to the clock.[11]

---

[8] *Ibid*., p. 230.

[9] Desgabets, *OPD* 4.132–133.

[10] Descartes (1649/1989, p. 21).

[11] Ricci (1953).

But we do have witness to the reaction of the natives from the Jesuits on the other side of the world, in fledgling New France. The Lilliputian's reaction was not entirely fictional, but rather, was an anthropological observation. If the story came from the Jesuits in New France, then perhaps Boyle mixed up the Americans with the Chinese, or changed the anecdote to give added emphasis to the example. Jean Brebeuf, the famous Jesuit missionary who was eventually tortured to death by the Iroquois, wrote to Father Le Jeune in "Kebec:"

> As to the clock, a thousand things are said of it. They all think it is some living thing, for they cannot imagine how it sounds of itself; and when it is going to strike, they look to see if we are all there, and if some one has not hidden, in order to shake it. … They think it hears, especially when, for a joke, one of our Frenchmen calls out at the last stroke of the hammer "that's enough" and then it immediately becomes silent. They call it the Captain of the day. When it strikes they say it is speaking; and they ask when they come to see us how many times the Captain has spoken. They ask us about its food; they remain a whole hour, and sometimes several, in order to be able to hear it speak.[12]

It is likely that Father Brebeauf's anecdote is the primary cause of Swift's joke,[13] and Boyle's anecdote, for their similarity is remarkable. The Huron also note the controlling effect of the clock upon the Frenchman in their perceptive description of the clock as "Captain of the day". Its self-motion, according to Brebeauf, leads them to think of it as alive, and its power over the Frenchmen, to it being a captain.

These stories are fascinating for how the epistemology of wonder, mechanism and life were being tied together by mechanists and the popular imagination. They are fascinating in how rumors of the New World, whether it is the re-discovered Chinese world or that of the Americas, were being brought to bear to make the epistemology of mechanism more palatable. However, I don't wish to exaggerate their role since they are mere anecdotes—evocative that they are. To understand its significance and the role wonder plays in these anecdotes, we should turn to Descartes' and Malebranche's analysis of wonder. Wonder is utilized in these anecdotes, and it is the use and the misuse of wonder that concerns me most. I shall return to these anecdotes and show how wonder contributes to their rhetorical force. Cases like these help reveal how wonder can play only a limited and momentary role in the life of inquiry, but its limited role and its misuse help explain the persistence of ill-conceived science.

---

[12] *Black Gowns and Red Skins; Adventures and Travels of the Early Jesuit Missionaries in North America (1610–1791)* Selected and Edited by Edna Kenton, 1954 Longmans, Green & Co. 1st edition published 1926, as "Jesuit Relations and Allied Documents (1610–1791).

[13] Although Thomas Harriot brought "spring clocks that seeme to goe of themselves" and showed them to the Virginians, he records only that their admiration for the Europeans increased.

2 Wonder and Passions of the Soul

Descartes' *Passions of the Soul* is an ethical book written from the natural philosopher's point of view. Descartes follows his contemporaries in combining physiological reflections with what appear to be moral recommendations for the management of one's well-being. Thus, we are to hear simultaneously a natural history of the passions and a discussion on their appropriate and inappropriate uses. The passions themselves are constituted by ethical judgments; that is, the external objects that eventually lead to a movement in the soul are referred to how they will be good or bad for the body. The causation of passion from an external object is itself not merely dependent on the "diversities"[14] or properties of the objects, but only in proportion to the many ways that the object can harm or profit us. The passions represent or respond to the objects as good or bad *for ourselves*.[15]

The use of the passions, generally speaking, is to "dispose the soul to will the things that nature tells us are useful", and Descartes here supposes that we may begin with a broadly optimistic, teleological attitude towards the passions and the body machine. The body–machine is often compared to the watch;[16] so, in each example of Descartes' use of the body–machine metaphor, the teleological commitment is found implicitly.[17] The body–machine can sustain its vegetative functions without the need of the will, but it is impossible to doubt that these functions have the sustaining of the person's life and well-being as their end. Descartes doesn't trumpet this side of his physiology, for his official view regarding final causes in natural philosophy requires that we remain silent about them, and that goal-directed behavior was reductively explainable.[18] But Malebranche hasn't the exact same scruples, writing buoyantly:

> Nothing is more marvelous than this arrangement of our passions and this disposition of our bodies with regard to objects surrounding us. All that mechanically takes place is worthy of the wisdom of Him who has created us. And, as God has made us capable of all

[14] Descartes (1649/1989, p. 51, article 52).

[15] Descartes' discussion of the manner in which our passions refer to objects is not clear to me. At times it seems as if the cause of, say, wonder, is the greatness and novelty of the object, but these are not clearly the content of a state of wonder at an object. When one wonders at Big things, one has esteem for that object, but it isn't clear exactly what the content of this wonder is: is it "wow that's a big novel object" because it was caused by a big novel object? See Lilli Alanen (2003b).

[16] Descartes (1649/1989, article 16).

[17] That is, like Boyle in *Disquisition About Final Causes* (1687), I think that Descartes cannot escape this commitment to the body being designed for our benefit. Descartes doesn't invoke the natural ends of the organs to explain their behavior, but steadfastly attempts mechanistic accounts of the actions. But such accounts are not inconsistent with the teleological assumption, according to Boyle. Furthermore, Descartes' attempted reduction of goal-directed motion in terms of the properties of matter was considered to be hopelessly inadequate.

[18] But see Simmons (2001) for a teleological reading of Descartes' theory of sensations.

> the passions that move us mainly in order to link us to all sensible things for the preservation of society and of our sensible being, His plan is so faithfully carried out in the construction of His work that we cannot fail to wonder at its construction and design.[19]

The passions form a halfway house between mind and matter and perhaps allow us to better conceive the union of soul and body. The passions are relational in that they are "perceptions or sensations or excitations of the soul which are referred to it in particular and which are caused, maintained and strengthened by some movements of the spirits."[20]

Passions are something like secondary qualities, and Malebranche puts emphasis on this analogy, worrying that the same errors made by the uncritical use of our perception can occur with our passions.[21]

As a passion, wonder exhibits both a passive and active side. "Wonder is a surprise of the soul, which makes it tend to consider attentively those objects which seem to it rare and extraordinary,"[22] writes Descartes. The passion is aroused by the objects when they are judged to be novel, but it is also an action directing us towards an inquiry into the object. Wonder has uses in addition to drawing our attention to objects, but these uses, good or bad, will not be essential to the passion as such. So it is the dynamic process of surprise leading to attention, which should be understood as wonder. "Wonder" is here the translation of the French and Latin *L'admiration* and *admiratio* respectively, and there is thus an association with admiration, a mode of being drawn towards an object.[23] Indeed, wonder was not an uncommon attitude to have towards God and his works, where admiration would find a natural place.

We find Descartes' contemporaries using wonder to ill effect and I believe Descartes can be seen as correcting their excesses. Looking carefully at the typical phenomenology of wonder we can note how it best suits a non-Cartesian and broadly Aristotelian ontology, and that means that it will have serious limitations for Descartes. Thus, wonder has its uses but it also has a fairly circumscribed role and must be curbed and controlled like a passion. It is the image of the inquirer and the proper place of wonder and curiosity within science that is at issue.

---

[19] Malebranche (1674/1997, p. 277). Malebranche differs from Descartes here—having a more broadly social account of the significance of the passions. For Malebranche the passions are very public phenomena and give rise to emotional contagion through our visible, nearly reflexive, bodily expressions. This public expression of the passions is, as one might expect, both good and bad, depending on the example and what is being conveyed.

[20] *Ibid*., p. 34.

[21] Malebranche: "…that sensible pleasure stands to our good as our sensations stand to the truth, and that just as our senses can deceive us with regard to the truth, our passions deceive us with regard to the good." (1674/1997, p. 362)

[22] Descartes (1649/1989, p. 56, article 70).

[23] See Eamon (1993) for discussion.

But wonder, being different from the other emotions, is key to living well and regulating our passions. This is because self-esteem and generosity are species of wonder. Wonder at ourselves is legitimate, and will be greatly beneficial, for it reveals us as we truly are—free. But this realization, consistent with wonder's nature, cannot involve a serious science. Indeed, our freedom is a proper object of self-regarding wonder; therefore, our freedom must itself be forever inexplicable and marvelous.

Descartes and Malebranche share a similar analysis of the causes of wonder. Both take wonder to be caused by "rare and extraordinary" objects,[24] that is, by objects that are novel to the understanding. Both also agree that wonder becomes esteem when the object wondered at is great, and when the object is small, our wonder becomes scorn.[25] And when the esteem is for oneself, this becomes magnanimity or pride, and humility or servility. Finally, when the esteem or scorn is directed towards a free agent, esteem gives rise to veneration, disdain from scorn.

Esteem is especially important, for directed at ourselves, we find that the one thing we may legitimately esteem is our freewill. From a proper appreciation of our freewill we become generous—understanding that who we are is within our control, and those things beyond our will are of little importance. Generosity is Descartes' ideal or moral for *The Passions.* Acknowledging our freewill, our self-esteem is legitimated and keeping in mind the assurance that the Will always has control over our passions, even if it is indirect, we are given the hope that the cultivation of generosity can lead us out of the problems of the soul. The control of the passions, which is essential to achieving happiness, can stem from the cultivation of generosity.

What is curious about this analysis, however, is that Descartes and Malebranche deny that wonder involves a judgment that the object is good or bad *for us*. It is difficult for us to understand how esteem and scorn can be morally neutral for they appear to be "pro" and "con" attitudes; esteem represents the object as good and scorn as bad, and certainly our 21st century phenomenology of scorn appears to involve a negative judgment.26 Susan James makes this complaint against Descartes arguing that:

> Esteem for grandeur can amount to no more than the recognition that someone possesses a lot of something by the going standards, whether power, jewels, learning, or sheer bulk. Equally, it can consist in appreciation of their non-moral qualities, as when someone is esteemed for their exquisite clothes. But because both of these kinds of assessment shade swiftly into evaluation of the good and harm someone may do us, there is no firm boundary separating them from their moral counterpart. It is difficult to exclude moral

[24] Descartes (1649/1989, p. 56, article 69; p. 59, article 75).

[25] Malebranche (1674/1997, p. 376).

[26] James (1997, p. 170) for discussion.

> evaluation from esteem and contempt. And it is therefore difficult to sustain Descartes' view that the latter are distinct passions, but are not concerned with our good and harm.[27]

But according to Descartes, wonder is the first passion and a primitive passion. It is primitive in that it is not a function of another passion or passions: it is not a species of another passion, nor is it composed of other passions. It is deemed to be the first passion because, unlike the other passions, wonder doesn't involve a judgment of how the object is good or bad *for us*. The object is not referred *to us*, although some judgment regarding the object's novelty is implied. It is thus not so obviously a moral passion, but an intellectual one. Wonder engages us towards understanding an object or draws us towards the first image of the object, while we are simultaneously being surprised by the object. Wonder is like a passion however, for like the passions it has to be governed by the will, as all passions must. Second, wonder can be considered as a passion, for we find the spirits actively re-enforce and intensify the passion, although in a considerably different physiological manner than the other passions. Since wonder is relevantly similar to other passions, despite its non-moral focus, the need to articulate wonder's healthy and unhealthy uses follows.

But it is likely that it is also called "first" due to the immediate neuro-physiological causes of wonder, which are closer to the soul than are the animal spirits. Descartes writes that the passion is caused by "the impression in one's brain that represents the object as rare"[28] and *then* it is caused by the motion of spirits that *reinforce* the impression of novelty, while holding the body's limbs in position. But although spirits are involved in the reinforcement of novelty, the spirits in the blood and heart are not the cause of this passion. Descartes' remarkable reason for believing this physiological claim is that wonder lacks good and evil as its intentional object or judgment.[29] He writes:

> …(I)t is not observed to be accompanied…by any change taking place in the heart and in the blood. The reason for this is that, not having good and evil as its object, but only knowledge of the thing wondered at, it has no relation to the heart and blood, which all the good of the body depends on, but only to the brain, where the organs of the senses are that contribute to this knowledge.[30]

Now, putting aside Descartes' apparent teleological reasoning, it is clear that Descartes is going out of his way to distinguish wonder from the other passions and to conceive of wonder as a morally neutral reaction to an object. Esteem then, as James notes, can involve no more than a recognition of the object and its properties, so that we could have esteem for someone due to their wealth or power in the sense that we are recognizing that they have such wealth and power. But

---

[27] James (1997, p. 170).

[28] Descartes (1649/1989, p. 57).

[29] *Ibid*., p. 57, article 71.

[30] *Ibid*., pp. 57–58.

James thinks that esteem shades into moral evaluation and that "there is no firm boundary separating them (having esteem towards things) from their moral counterpart."[31] But I don't see why she thinks this slippery slope argument is a fair criticism, and it is certainly not something that Descartes is committed to.

In defense of Descartes, I think we can evaluate and attribute value to something without being committed to a *moral* evaluation. Can I not have esteem for the beautiful, without judging that the beautiful person, or the beauty, is better *for me*? Descartes notes that "esteem" and "scorn" are often names given to a person's *passionless* opinion of the worth of an object, but because a passion often seems to arise on such occasions Descartes also labels these "esteem" and "scorn."[32] But Descartes is self-consciously extending the use of these terms to apply to what he considers to be previously unnamed passions. Aware of the extension he writes: "And Esteem, *insofar as it is a passion,* is an inclination the soul has to represent to itself the worth of the thing esteemed."[33] Descartes is certainly making every attempt to show that esteem and scorn are species of wonder, and hence do not imply anything regarding how the objects of these passions affect us. Descartes appears to allow "esteem" to be the name of a passion because he notices that the animal spirits reinforce or strengthen the idea of the thing's worth. But unless we eschew the distinction between something being worthy or good in itself and something being good or worthy for me, Descartes is surely not faced with the slippery slope that James suggests. Wonder, and hence esteem, focuses our minds on the object itself, or the appearances of the object, and makes a judgment about the object without comparing or referring that object to our possible benefit or harm. I can't see why the existence of such a non-morally loaded "passion" is to be doubted, *pace* James. To the extent we do think that esteem involves some form of positive evaluation we can rescue Descartes from contradiction if we recognize that positive evaluation need not be moral evaluation—i.e., it need not be evaluation with regard to how my well-being is affected by the object.

3 Wonder and Inquiry, Rainbows and Mechanism

One consequence of wonder implying ignorance is that wonder is something that ideally will disappear. The Cartesian will cease to feel wonder at objects just as those objects become explained and accounted for. Understanding an object is incompatible with surprise at the object, so once we have given an account we will no longer be in a state of wonder. Were we to explain the world, all of it, we would

[31] James (1997, p. 170).

[32] Descartes (1649/1989, article 149).

[33] Emphasis added.

feel no wonder at all. The ideal of science could be put this way: to eliminate our wonder at nature with knowledge. The world itself is not wonderful at all.

Descartes exhibits this attitude in his early work on the rainbow.[34] But the idea that wonder is eliminated by proper inquiry and knowledge is not unique to Descartes, but is found in most 17th century writers on the topic, and with Aristotle. When Descartes published *Discourse de la Methode* (1637) he also included as an appendix his essay *Les Meteores*. Writing to Mersenne, Descartes remarks that paying attention to *Les Meteores* and his work on the rainbow, one would be able to understand his work on method better.[35]

The rainbow was certainly something of a wonder with its splendid appearance.[36] In its obvious connection with water and light it simply advertised its possession of hidden secrets. Descartes had read Jean Leurechon's *Recreations Mathematiques* (1624).[37] Leurechon's book consisted of somewhat random observations, "secrets and experiments," on almost everything useful, such as "Arithmetick, Music, Opticks, Water-works, fire-works, Mechanicks." His book was not unusual in its attempt to compile useful and entertaining "facts." In problem XLVI (44 of the English version) Leurochon writes: "The rainbow is a thing admirable in the world, which ravisheth often the Eyes and Spirits of men…." Leurochon offers advice regarding how to see the rainbow effect by spraying water from one's mouth. Leurochon's intention is to show us how to reproduce the phenomena in question, but not to give any significant explanation. Descartes would have none of that and he set forth to offer serious explanations of the rainbow and the clouds, in accordance allegedly, with his *Discourse on Method*. Although the rainbow is his target Descartes writes the following on the clouds:

> It is natural for us to have more admiration for things that are above us than for those that are on the same level or below us. And although the clouds are scarcely higher than the summits of some mountains… because we must turn our eyes towards the sky to look at them, we imagine them to be so high that poets and painters even make them into God's throne, and picture Him there … This leads me to hope that if I explain the nature of the clouds here, in such a way that we will no longer wonder at anything that we see of them, or that descends from them, we will find it easy to believe that it is likewise possible to find the causes of everything that is most admirable above the earth.[38]

Our wonder will be relieved when we have an explanation of the phenomena, and the success of such a project illustrated in the three appendices to the *Discourse* gives us confidence that we may also apply the same principle to the superlunary world. What is significant is Descartes' clear commitment to explaining nature in a way that requires us to cease wondering. This is a significant attitude—there

---

[34] See Armogathe (2000), Boyer (1959).

[35] See Gaukroger (1995, pp. 217–219), Descartes to Mersenne 13/11/1629, AT I.70.

[36] See *Ecclesiastes* 4.3; Descartes AT VI.325.

[37] To Mersenne, April, 1634, AT I.285.

[38] Descartes, AT I.25.

were many thinkers who paid more attention to reproducing effects and the wonder those effects cause, than to serious explanation of the effects. Descartes writes: "…although it is good to be born with some inclination to this passion, since it disposes us to the acquisition of the sciences, we should still try afterwards to emancipate ourselves from it as much as possible."[39] Wonder encourages us to inquire after the object, but there are good and bad uses of wonder. It is to these that I now turn.

4 Good and Bad Uses of Wonder

What, then, is wonder good for? It will, of necessity, disappear when we have offered proper explanations of the phenomena. But wonder has a dual role as we have already noted. In the form of self-esteem and generosity wonder at ourselves is crucial for the proper recognition of our limits and our powers—that of freewill. But when directed towards the objects of a potential science, wonder is important for the role it plays in getting us to pay attention to the object and to remember the object. Apparently, even though we find the object fascinating in its novelty, this is insufficient for us to remember that object unless we have some passion acting on the brain, or, we are turning our understanding towards the object.[40]

The "surprise" in the soul and the bodily fixation which it implies, can be excessive—we can have too much wonder, resulting in the body becoming immobile and the mind unable to consider anything but "the first face presented"[41] of the object. This is our state when astonished by an object. We are like a rabbit caught in the spot-light. Astonishment is always bad, because it arrests both the body and the mind. The best remedy for astonishment—that is, how to avoid having it, rather than how to get out of it once one is gripped by the passion—is to "apply oneself to the consideration of all those which may seem most rare and unusual."[42] But there are several other problems with wonder also. First, astonishment can become habitual, or lead to habitual wondering without the real work of understanding following after, as it should. This habit can lead people to search for wonder for its own sake—to simply wonder without coming to have knowledge. And as a result of this kind of habit our wonder can give off false positives and false negatives. Wonder can be directed at objects of no importance whatsoever, and wonder can fail to be triggered where it should.

---

[39] Descartes (1649/1989, p. 60, article 67).

[40] Descartes (1649/1989, p. 59, article 75).

[41] *Ibid*., p. 58, article 73.

[42] *Ibid*., p. 58, article 73.

The latter problem—that of delivering false negatives—is illustrated by Malebranche in a remarkable passage in which Descartes himself becomes the object one ought to have wonder for. Malebranche complains about the obsession and reverence his contemporaries had for the ancients and offers an anecdote about a man who met Descartes: "This fellow had a low opinion of Descartes' philosophy because he spoke with the man for a few minutes and failed to notice anything about him that smacked of the wonderfully great appearance which warms the imagination."[43]

It is noteworthy that wonder is often concerned with the appearance of the object, and this is one of the problems with wonder. Descartes was considered a small frail man, certainly not a big guy, but Malebranche is saying that Descartes didn't have the appearance and demeanor that would produce wonder, although Descartes deserves such wonder for being so brilliant. Wonder can dwell on the "first image of the object"[44] without going further into a proper inquiry. Wonder is typically a reaction to an object presented to the senses, so it can become fixated and misled by the mere appearance of the object.

Wonder is a step towards science and, as such, will eliminate itself for it tends to focus us upon the appearances of the object rather than a proper understanding of the object. But Descartes and Malebranche's cautions regarding wonder can be best understood by considering their historical and intellectual context. Descartes had reason to make these cautions against wonder, for his contemporaries often had different attitudes. Malebranche's discussion of wonder is intertwined with his criticism of Virtuosi who spent much of their time on wonderful antiquities and manufacturing splendid effects. For example, Jesuit thinker Anathasius Kircher (1602–1680), to take the most famous, was well-known throughout Europe for his clocks, inventions and wonderful machines. In 1633 Kircher had advertised his "sunflower" clock whose seeds allegedly followed the motions of the sun when the flower was placed on a cork in water. But Kircher had a mixed reputation and was, by the time Malebranche wrote *The Search after Truth,* considered by most to be rather unreliable and close to fraudulent in the explanations of how his wonderful machines worked. But in 1633, when wonderful new discoveries seemed to be overly abundant, one couldn't be quick to dismiss anything. Kircher's "sunflower clock" turned out to be a trick with magnets. As noted above Leurechon was concerned with reproduction of wonderful affects without offering up any serious explanations. Let us return to the Jesuits over in New France.

The Jesuits were explicit in their use of wonder with regard to the Huron.

> Speaking of their expressions of admiration, I might here set down several on the subject of the loadstone, into which they looked to see if there was some paste; and of a glass with eleven facets, which represented a single object many times, of a little phial in which

[43] Malebranche (1674/1997, p. 383).

[44] Descartes (1649/1989, p. 61, article 78).

> a flea appears as large as a beetle; of the prism, of the joiners tools; but above all, of the writing; for they could not conceive how, what one of us, being in the village, had said to them, and put down at the same time in writing, another, who meantime was in a house far away, could say readily on seeing the writing. I believe they have made a hundred trials of it. All this serves to gain their affections, and to render them more docile when we introduce the admirable and incomprehensible mysteries of our Faith; for the belief they have in our intelligence and capacity causes them to accept without reply what we say to them.[45]

The Hurons' wonder becomes admiration and allows them to be more susceptible to even further admirable mysteries. The Jesuits are clearly not concerned with relieving their subjects of their wonder by offering up explanations for the items displayed. Indeed, many of these wonders were lacking convincing explanations for the Europeans themselves, which reminds us that the experience of wonder at these mechanisms was also being felt by many Europeans.[46] But the Jesuit's use of wonder is clearly not as an aid to memory, as Descartes claims wonder to be useful for. Nor is wonder being used as a spur to inquiry regarding the underlying causes of the objects. Rather, wonder is used to make the subjects ready and docile for the acceptance of further mysteries. Although the spiritual goal is admirable, wonder is not being used as a tool for memory as Descartes advocates, nor is it the beginning of more rational inquiry. Descartes does not ever explicitly say that wonder is well-used as a tool to encourage devotion, although he allows that wonder at God, not nature, is appropriate.

Giambattista Della Porta (1565–1615) also has a curious attitude towards the use of wonder and, like the Jesuits, Kircher and Leurechon, Descartes had read Della Porta's well-known treatise *On Natural Magic*. In a section entitled "…what manner of man a magician ought to be," Della Porta notes that philosophy is required for knowledge of the effect of the four elements but that workmanship is essential to the enterprise:

> He must be a skillful workman, both by natural gifts, and also by the practice of his own hands: for knowledge without practice and workmanship. And practice without knowledge, are nothing worth; these are so linked together, that the one without the other is but vain, and to no purpose.[47]

But having said that, after having noted how one must be diligent in one's preparations and how the ignorant may take the phenomena to be "haphazard" rather than following of necessity from underlying causes, Della Porta reminds us how one might make the phenomena *remain* wonderful:

> If you would have your works appear more wonderful, you must not let the cause be known; for he that knows the causes of a thing done, doth not so admire the doing of it,

---

[45] op. cit. *Black Gowns and Red Skins* pp. 110–111.

[46] See, for example, Robert Hooke's magnificent book *Micrographia* (1665).

[47] Della Porta (1669, p. 3).

> and nothing is accounted unusual and rare, but only so far forth as the causes thereof are not known.[48]

Della Porta is still beholden to a tradition of secret crafts and secret recipes, so his motives are still tied in with the economics of knowledge and craft.[49] Wonders are to be investigated and Della Porta is keen to trumpet his revelation of the secrets; but he is not unhappy with keeping some people enthralled in their wonder. In the English translation by Van Etten of Leurochon's *Recreationes Mathematique*, we get a similar sentiment being aired. One of the intentions of the book is

> To give a greater grace to the practice of these Things, they ought to be concealed as much as they may, in the subtilty of the way; for that which doth ravish the Spirits is, An Admirable Effect, whose cause is Unknown; which if it were discovered, half the pleasure is lost: therefore all the finess consists in the dexterity of the act, concealing the means, and changing often the stream.

What both of these examples reveal is just how the proper use of wonder, in this case as a temporary stepping stone towards real inquiry, was not shared by all in the 17th and late 16th centuries. Indeed, the mechanists had been almost too clever, in a sense, for their own machines had become objects of wonder. People were being entertained by the wonder they experienced from their own culture's machines and there was money and reputation to be made in such activity.[50] A cultural problem was being addressed. The past continued to have its influence through those who wished to keep their secrets to themselves and their initiated few, and hence wonder was all these folk would allow one to have. Second, the developments in mechanistic technology, the talking fountains and clocks for example, were producing opportunities for charlatans and showmen to entertain their patrons by encouraging their wonder towards the secrets of the mechanism. But these attitudes and practices are distractions to science.

Wonder is also problematic in that it is typically a response to a perception of the object. Wonder can fixate upon the appearances of objects. When wonder fixes our attention it typically fixes our attention on visible properties of the object—hence it failed to be triggered when directed at Descartes' plain and modest appearance. This suggests a further worry: that wonder, although helping us towards inquiry could mislead us into taking the appearances as overly important. And again, there were plenty of people who fell for that. Della Porta takes the appearances to be important in many ways: his experimentalism is always directed towards producing observable results from the recombination of substances, based on their visible qualities. Della Porta's Aristotelianism is part of the issue here. But a science based on the secondary qualities, whether experimental or not, will never succeed. Wonder, however, focuses our attention on these secondary qualities

---

[48] *Ibid*., p. 4.

[49] See Eamon (1993).

[50] See Stafford (1994).

and could thus distract one from a proper Cartesian science. Della Porta also took the appearances to be ontologically significant in that he was a supporter of the doctrine of signatures, holding that the appearances of certain simples hold clues to their utility. Such a science would indeed find wonder appropriate, as it would focus the mind on the qualities and appearances on the novel objects. But such a science would not be a Cartesian science.

Concluding Remarks

We have seen that Descartes' notion of wonder plays two roles in *The Passions of the Soul*. The first is part of the key to happiness. Our wonder at ourselves reveals our freedom which gives us the key to the control of the passions and the proper attitude towards fortune.[51] In its second role wonder is important as a way to encourage us to inquire and further proper science. But its role is limited, for it is a passion after all, and it can seriously mislead us, by becoming its own end, or in encouraging a science of the appearances.

Let us return to the Huron and the clock. The Huron wondered and were led to admire the Jesuits (as the Jesuits flattered themselves to believe). The Huron took the clock to be alive, which of course is an error. The way wonder functions here is this. The clock moved and spoke on its own, and wondering at the clock's appearance one doesn't know the cause of this self-motion. But if the cause of self-motion is unknown then it would be correct to conjecture that the object is alive; after all, the principle behind the self-motion is occult and self-motion will require an active principle or soul. The Mechanist's and the Cartesian's conceit is that this natural inference based on wonder at self-motion is a mistake. Wonder can mislead the naive. The natives like the naive Aristotelian has let their wonder mislead them and Swift's Lilliputians follow suit.

The Jesuits in New France, more concerned with religious education and conversion than with explanation, show few scruples in using this wonder to ready their Huron hosts for Christianity. Descartes remains silent on the religious use of wonder at nature (or other machines), for he cannot maintain wonder for religious purposes while simultaneously sustaining his commitment to offering complete physical explanations of nature.

Indeed, in some 17th century thinkers the rejection of wonder at nature was an atheist conceit, so it is understandable that Descartes is reticent in this matter. Samuel Butler's (1612–1680) witty and biting satires convey the temper of many of the period. Under the entry for "An Atheist" Butler writes:

[51] See Brown (2006).

> Nothing but Ignorance can produce a Confidence bold enough to determine of the first Cause; for all the inferior Works of Nature are Objects more fit for our Wonder, than Curiosity; and she conceals the Truth of Things, that lye under our View, from us, to discourage us from attempting those, that are more remote. He commits a great an Error in making Nature (which is nothing but the Order and Method, by which all Causes and Effects in the World are governed) to be the first Cause, as if he should suppose the Laws, by which a Prince governs, to be the Prince himself.[52]

Butler appears to favor wonder over curiosity and piety over knowledge. The Cartesian's attitude to wonder, therefore, cannot be taken on by everyone. Wonder lived on as the rhetorical and emotional force underlying the teleological arguments for God's existence and for the existence of teleology generally. Such teleological reasoning was explicitly rejected by Cartesian science. Despite Descartes' ambitions philosophers *could not* relieve their wonder with an account of the mechanism of life—that would not occur until the 20th century—but if wonder persisted, then mechanism was constrained, and the intelligent hand of God, or his lesser vehicles, the vegetative soul or the world-soul, was once again required to save the appearances. But how wonder functions in the teleological argument is a tale for elsewhere.

[52] Butler (1970, p. 163) (emphasis in original).

Desgabets: Rationalist or Cartesian Empiricist?[1]

Sean Allen-Hermanson

My interest in this project owes to curiosity about the comparison sometimes made between the "first" and "second" cognitive revolutions. The second revolution is the one that began in the mid-20th century and which continues today.[2] It is characterized by the replacement of prevailing frameworks (e.g., Behaviourism) for exploring memory, learning, language, and thinking, by new techniques and theories inspired by the metaphor of mind as an information processor, like a computer, that solves problems by applying logical transformations to internal symbols. At risk of oversimplifying a complex history, the more recent cognitive revolution represents a shift in philosophical attitudes back towards a broadly Rationalist outlook when it comes to understanding higher mental functions, i.e., one that de-emphasizes sensory experience, training, and individual history, in favour of innate mechanisms, not dependent on sense perception, and orientated on the development of the species as a whole.[3] The source of this change can be traced to similar change in approaches to the mind in the 17th and 18th centuries. Chomsky (2005) has urged continued study of this first cognitive revolution—especially the contributions of the Rationalists—in order to recover insights still useful to the second. Although I fear the sin of over-enthusiasm, I can at least agree that a re-examination of how we got where we are can sometimes lead in surprising new directions.

With these varying degrees of ambition in mind, I turn to examine a little-known philosopher from that earlier period, Robert Desgabets (1610–1678), who has idiosyncratic views on several areas of continuing interest, including the nature of representation, modality, and time. But I will restrict myself to a discussion of what he has to say about the role of sensation in the formation of ideas. Desgabets was a Cartesian and contemporary of Descartes, and commands attention for his seemingly unusual blending of Cartesianism and empiricism.[4]

---

[1] I thank Patricia Easton, Jon Miller, participants of Topics in Early Modern Philosophy of Mind, held at Queen's University in 2006, Louis-Philippe Hodgson, and an anonymous reviewer for their kind advice, patience, and helpful criticisms.

[2] Some pinpoint its birthdate as September 11, 1956 and the conference held at MIT where seminal research in linguistics (Chomsky), psychology (Miller), and computing (Newell and Simon) were presented (see Gardner 1985, p. 28).

[3] This is just the big picture. Certainly other emerging frameworks, such as the connectionist alternative, and dynamical systems theory, are trying to pull things in other directions.

[4] I have adopted the following notation system for citations of Desgabets's *Supplément à la philosophie de Monsieur Descartes*: Arabic numerals denote parts 1 or 2; capital roman numerals denote the chapter; and lower case roman numerals denote section number, e.g. S2.I.iii denotes

J. Miller (ed.), *Topics in Early Modern Philosophy of Mind*, Studies in the History of Philosophy of Mind 9, 57–84.

Allow me to immediately acknowledge that the terms "empiricist" and "rationalist" are blunt instruments, and should not be carelessly employed.[5] Prototypical philosophers from each side often overlap in their fundamental assumptions, methods, and agendas; philosophers on the same side can likewise differ greatly.[6] A variety of criteria might be used to make this distinction with precision; I will not pretend that there is one way satisfying to all. However, I will assume that the distinction can be interesting, and useful. For my purposes, empiricism will be taken as the view that certain ideas, or concepts, are only gained through sense experience. Rationalism denies this, maintaining instead that the ideas are acquired in some other way: perhaps they are rationally intuited, or deductively constructed, or are somehow already present in the mind. Concept empiricism should not be confused with "knowledge empiricism" which will be taken to assert that certain truths or propositions can be justified and known only through the operation of sense experience. Empiricisms of differing strengths result from expanding or contracting the extension of the word "certain" in the previous sentence: the domain of one's empiricism is negotiable (e.g. perhaps including mathematics, morals, or metaphysics).[7] Having said that, let it be agreed much turns on the specific role of sense experience—depending on how it is defined, "rationalism" is compatible with saying that ideas are "gained," "acquired," or "constructed" from sensation. For example, one could accept innate ideas while maintaining that they depend on sense experience to trigger their occurrence. This is not what I have in mind. Content empiricism is not a claim about the causes of ideas, it is rather about the nature of representational content, specifically, that the content of ideas is borrowed (i.e. identical to) the content of sensory impressions, perhaps modified by such operations as combination, amplification, and diminishment. Hence, the claim in play here is whether Desgabets accepts concept empiricism.

Desgabets participated in important controversies of the day, including the question of atomism versus the infinite divisibility of matter,[8] and the nature of the

---

the third section, of the first chapter, of part two. '*CdC*' refers to Desgabets's *Critique de la Critique de la Recherche de la Vérité*. These are contained in Desgabets's *Oeuvres philosophiques inédites* which will be referred to as '*RD*.'

[5] See Kenny's introduction to his 1986.

[6] For example, Descartes, Leibniz and Spinoza are simplistically, and incorrectly, seen as advancing phases in an unified research program.

[7] Obviously concepts are prerequisites for knowledge, but concept empiricism and knowledge empiricism are logically independent, e.g. one could maintain that a certain class of concepts are acquired experientially, but warranted belief employing them is not, or vice versa: perhaps knowledge is experiential, but concepts are not.

[8] Easton (2006, p. 2) reports that in a letter to Clerselier about Cordemoy, Desgabets complains about the tension between the Cartesian and anti-Cartesian elements of Cordemoy's *Discernement du corps et de l'ame*. I haven't seen the letter, but its content might be relevant to the topic of this paper.

Eucharist.[9] His *Critique de la critique de la Recherche de la vérité* was an inapposite defense of Malebranche against Foucher, though not well received, most pointedly by Malebranche himself. Watson wryly observes that "it was as dangerous to defend Malebranche as to attack him."[10] This incident led to the unfortunate epithet: "The disciple of Malebranche who understands nothing of Malebranche"[11] and goes some way towards explaining Desgabets's obscurity.

His *Supplément à la philosophie de M. Descartes*, unpublished in his lifetime, is a defense and elaboration of many key Cartesian themes, but also takes issue with a number of others including pure intellection, the existence of innate ideas, and the concept of objective reality.[12] This work also discusses intentionality[13] and Descartes's doctrine of the eternal truths. The *Supplement* is most importantly a sustained examination of the nature of ideas, and in this chapter I will use it to address a dispute about Desgabets's alleged empiricism.

Easton and Lennon are recent sources for the view that he is a "Cartesian empiricist." As Easton says, "heretically to some, he strongly rejected the rationalist epistemology which often dominates in Descartes, and argued that Descartes's own principles favour a form of empiricism."[14] According to Easton, Desgabets held that "all (true) knowledge depends on the senses, and hence on our perception of…sensible qualities and objects."[15] He believed that "[t]he soul must always be in commerce with the senses, and…our thoughts depend on the corporeal traces in

---

[9] In the anonymously published *Considérations sur l'état présent de la controverse touchant le T. S. Sacrement de l'autel* Desgabets was the first to openly propose that the body of Christ is literally present in the host. This led to a backlash against Cartesianism as this was thought to be incompatible with official Church Dogma.

[10] Watson (1987, p. 256).

[11] Guéroult (1987, p. 254). "Il me semble que ceux qui se mê lent de deffendre ou de combattre les autres, doivent lire leurs Ouvrages aver quelque soin, afin d'en bien sçavoir les sentimens" (Malebranche quoted in Watson 1987, p. 161).

[12] Desgabets accepted mind–body dualism, a substance-mode ontology, mind–body union, psycho-physical interaction, and the essential and contrasting natures of *res cogitans* and *res extensa* and a variety of other Cartesian themes and assumptions. Against Descartes, he held that the human "reasonable soul" is a third simple substance that "emerges" from the more fundamental union of the mental and the physical.

[13] Desgabets argued that the nature of intentionality could be used to construct a proof of the external world: he contends that one cannot think of what does not exist—thinking consists in the forming of "simple conceptions" of immutable substances which are the essences of modal things (Cook 2002). His reasoning here seems to anticipate an implication of Meaning Externalism—if intentionality consists in a relation (e.g., causal) between a representation and its object, then the intentionality of thought defeats skepticism about the external world (cf. Seager 1999).

[14] Easton (2001, p. 3).

[15] Easton (2001, p. 6).

the brain [though they are not identical to brain processes]."[16] This is most evident from the fact that thoughts endure for a finite time, and are succeeded by new thoughts. For Desgabets, this implies that they depend on *motion*, for time is essentially a measure of change in parts of matter[17]—and, as noted by Easton, ideas about motion can only be imparted by the sense organs.[18] This also implies that there can't be "pure intellection" in the sense of thought entirely divorced from motion, since this would not permit any distinctions between the beginning, end, duration, and succession of thoughts.[19] Their assessment of Desgabets's epistemology is echoed by many others, including Cousin, Bouillier, Rodis-Lewis, Ayers, and Lennon.[20] Schmaltz, with hesitation, and in a qualified way, also describes Desgabets as a Cartesian empiricist along the lines of Regius and Rohault.[21]

An interesting challenge to this long-standing agreement has recently been offered by Cook.[22] Cook reconsiders three aspects of Desgabets's thinking that have led others to conclude he is an empiricist: Desgabets attacks what other rationalists have said about pure intellection; he seems to endorse Locke's division of the mind into internal and external senses, stressing their central role in all forms of mental activity; and he endorses the empiricist doctrine that there is nothing in the mind save for what was previously in the senses. On this basis the case for his empiricism might seem highly plausible. However, Cook argues that, despite appearances, this evidence is superficial: Desgabets's is not endorsing concept empiricism, although he is asserting a strong causal dependence of the mental on the physical, situated within a broadly Cartesian, and especially Dualist, metaphysical framework. In short, Cook's point is that Desgabets is only saying all ideas are caused by motions in the body and the brain, and a rationalist dualism can certainly agree that all mental activity depends on, or is brought about by, brain processes.

I admit to finding parts of this alternative interpretation subtle and persuasive. Cook shows how the case for Desgabets's alleged empiricism is muddled when metaphysical and epistemological aspects of his theorizing are conflated. Although I think Desgabets was an empiricist of some sort, the demonstration of this needs to be sensitive to Cook's objections. So, allow me to consider the reasons against taking Desgabets to be a kind of empiricist more closely before turning to a

---

[16] Easton (2001, p. 8).

[17] *RD*, p. 299.

[18] See Part I, chapter 3 of the *Supplément*.

[19] Easton (2002, pp. 205–206).

[20] Cousin (1945), Bouillier (1868), Rodis-Lewis (1993, p. 423), Ayers (1998, pp. 1029–1030), Lennon (1993, p. 210), Lennon (1998, p. 353).

[21] Schmaltz (2002, p. 16).

[22] Cook (2008).

defense of the traditional reading. It will be shown that the case for empiricism is quite strong after all.

I begin with Desgabets's endorsement of the empiricist slogan *Nihil est in intellectu quin prius fuerit in sensu* ("nothing is in the intellect except what was previously in the senses," hereafter the *Nihil* principle). According to Cook, Desgabets does not agree that all ideas are formed or derived from sensory impressions in any robustly empiricist manner, and he even accepts Arnauld's and Descartes's view that none of our ideas exist as they do in the senses.[23] Desgabets actually modifies the *Nihil* principle to read "*from* the senses" (*a sensu*) in order to underscore that is only a claim about the causal necessity of the body, and denying that our "thoughts be similar to what happens in our senses"[24] Desgabets introduces this change in response to Materialist "Libertines" and Cartesian Dualists who mistakenly assume that the *Nihil* principle implies mind and body are not distinct; in this, both camps confuse strong causal dependence with identity. This mistake follows from the supposition that whatever comes from the senses must resemble physical parts of the nervous system. Desgabets means to correct this error by showing us that ideas can depend on the "senses," here understood broadly as "brain processes," not perceptions, or even the sense organs specifically, despite the metaphysical separation of extended and non-extended substances:

> [I]t agrees with the [Libertines] in recognizing that in all thought without exception there is something going on in the body, and it agrees with the [Dualist Cartesians], in that it holds that the soul, which thinks dependently on the body, is not at all the body, but that it is simply united to it …[25]

So, Cook argues, Desgabets only wishes to use the *Nihil* principle to make a point about the metaphysics of mind and body, not the epistemic conditions for knowledge, or for having ideas. The change to "from the senses" supposedly makes the maxim "much less empiricist … since a non-empiricist can surely say that the body and the senses cause our thoughts."[26] Even Descartes agrees to a limited dependence of the mental on the physical in that sense experience is "occasioned" by motions of the body. Occasional causation is perhaps most easily explained by distinguishing it from "transuent" causation. The latter conception assumes that something is transferred from cause to effect (such as motion, or pattern). Transuent causation also implies some sort of resemblance between cause and effect (i.e. whatever is transferred). Occasional causation is not like this: nothing is transferred, and there's no requirement that an effect resembles its cause.[27]

---

[23] Cook (2008, p. 506).

[24] S1.III.iv.

[25] S1.III.iii.

[26] Cook (2008, p. 512).

[27] Nadler (1994, pp. 38–39).

Arnauld and Descartes seem to have it that sensory mental contents ("sensible qualities" or the qualia of modern parlance) are not "transuently" derived from activity in the sense organs, but are somehow stored in the mind until the appropriate motion occasions their "release." The non-resemblance of our ideas to patterns in the nervous system is a crucial assumption. For if ideas do not resemble the things they represent, it seems to follow that their contents cannot be derived from sense experience or the external objects represented—then there would appear to be no reason to assert that thoughts are fundamentally *copies* of sensations; and if that is so, then the contents of our ideas must somehow be already present in the mind—for the specific character of thought has to come from somewhere. Desgabets maintains that all thought, including so-called "pure intellection," is occasioned by the body, but, not derived or copied from brain processes in any further sense.[28] Against Descartes, Desgabets insists that even the most "abstract speculations" imply "commerce with the senses."[29] But Cook suggests that Desgabets does not mean to reject Cartesian pure intellection for empiricist reasons—this is despite his sharing several of Descartes's assumptions, e.g. that it does not involve the formation of mental images[30] (though Cook concedes that Desgabets's assertion that "sense perception is sometimes imagination and sometimes pure intellection" does not elegantly fit this picture[31]). Schmaltz also takes the point that pure intellection includes sensations such as pleasure and pain to render Desgabets at least empiricish.[32]

According to Cook, what Desgabets really objects to in Descartes's view is just the claim that pure intellection can occur independently of the brain, and that it is properly pursued through the method of doubt.[33] Cook observes that these objections are compatible with rationalism. For Desgabets "species traced in the brain" are causes of all thoughts, including those obtained in pure intellection—but these brain states are not images, nor do they resemble mental contents.[34] The rejection of the method of doubt is likewise not in virtue of a special role for the senses, but because of Desgabets's belief that what he calls "simple conceivability" implies actuality. Pure intellection is contrasted with mental imagery where neural causal antecedents do resemble what he calls "sensible objects."

As for the distinction between internal and external senses, Cook insists that although Desgabets accepts it, for him the so-called internal senses consist in brain

---

[28] Cook (2008, p. 515).

[29] S1.III.iii.

[30] Cook (2008, p. 513).

[31] Cook (2008, p. 514).

[32] Schmaltz (2002, p. 16). Cf. *RD* 5, p. 192, *CdC*, pp. 91–104.

[33] Cook (2008, p. 514).

[34] Cook (2008, p. 514).

processes ("species traced in the brain") in a very general sense, not necessarily related to experience or perception.[35] It is important to note that the sense organs themselves are often referred to as the "external" senses by Desgabets; the "interior senses" are brain processes that serve as causes of inwardly oriented ideas, especially, I contend, feelings and interior perceptions.[36] Cook maintains that Desgabets only adapts empiricist jargon to a different purpose—when he says that all ideas depend on the "senses" it is only to emphasize the very close nature of the mind–body union. Still, I wonder. It seems like he could make do without employing so much empiricist-sounding talk. Knowing that this might encourage people to misread him, why bother? But as I say, Cook's crucial point is that he means to include brain processes *apart from activity of the sense organs* when he speaks of the "senses" (i.e., in his version of *Nihil*) and, especially, the "internal senses." Although he might seem to be talking like an empiricist, he is only stressing the dependence of our ideas on brain processes, and not claiming that our ideas are sensory in nature. For example, when he criticizes Descartes's rejection of the *Nihil* principle, it is on the grounds that Descartes denies the "reciprocal commerce" with the internal senses.[37] This goes against such commonplace observations as that sleep, bad health, age, and sensory impairment all influence the character of one's thoughts.[38]

Cook also acknowledges some apparent inconsistencies with his interpretation. There is a reference to Desgabets's suggestion that the internal senses might be somehow parasitic on the external senses.[39] If this was strictly true, then it seems sense organ activity is given a privileged place in the formation of ideas after all. Cook admits that Desgabets isn't always clear about whether it is the body, the brain, or just the external sense organs that our ideas depend upon: "when Desgabets says that all our ideas come from the senses he gives a picture on which they all come from the external senses," though, he adds, Desgabets usually stresses both.[40]

Easton's observation that all thoughts depend on motion might strengthen this objection. Ideas about motion can only be imparted by the *external* senses; but then it seems that something intrinsic to the character of our thoughts (i.e., that they begin and end, and so on) specifically depends on the operations of the sense organs. Then again, this isn't the same as saying that ideas are *wholly* derived from sense perception—this seems only to be true of certain aspects of thoughts (e.g. their finite duration—and further notice that for this to be relevant, it would have to be that a representational *content* is copied from some sense organ event);

[35] Cook (2008, p. 509), *CdC*, pp. 111–112, 129; S1.III.iii.

[36] S1.III.iii.

[37] S1.IV.i.

[38] S1.IV.iv.

[39] Cook (2008, p. 511).

[40] Cook (2008, p. 511).

one could take different views on this: perhaps a three second flurry of motion in the visual system corresponds to the representation of a three second event in the thought which follows, then again, not every property of a representation represents something (e.g. the duration of a movie), so likewise, perhaps the duration of a thought is not itself representational. Desgabets seems to be unclear on this point.

In summary, Cook addresses three issues concerning Desgabets's alleged empiricism: there is his endorsement of the *Nihil* principle, his emphasis on the internal and external senses, and his re-evaluation of the faculty of pure intellection. Cook argues that each of these provide weak support for an empiricist interpretation. He claims Desgabets's use of the *Nihil* principle is in a metaphysical, not epistemic, sense; he is only pointing out that mind–body dualism is compatible with a very strong causal dependence of the mental on the physical. Likewise, although the non-physical mind metaphysically depends on the "senses," he just means it depends on "brain processes" in general. This is, of course, compatible with the rationalist denial that the contents of our ideas or knowledge are somehow copied or otherwise derived from sense *experience*. Finally, pure intellection also causally depends on physical brain functions, contrary to standard Cartesian theorizing, but this is not to reduce it to a variety of perception either.

I agree with Cook that some may have conflated the dependence of ideas on physical causes in the nervous system with concept empiricism insofar as he denies that there is a *body*-independent pure intellect.[41] Easton says this, for instance, when she describes him as holding that "No ideas are innate since all ideas come from the senses in that they depend on the movement of our sensory organs for their formation."[42] Easton also makes a similar claim in her Stanford Encyclopedia entry where she writes that "our ideas … depend upon the operation of the senses." Easton and Lennon also jump from the claim that all thought (even "rapture, contemplation, and ecstasy") depends on the body, to the conclusion that Desgabets must be rejecting any distinction between ideas and sensations.[43] This criticism should be nuanced, however. These other authors are assuming he means ideas are *both* caused by physical processes, and constructed out of sense experience. Though I think the claim is correct, this will take a bit of work to demonstrate.

Before moving on, perhaps it is also helpful to keep in mind that even if Cook's criticisms are cogent, this does not establish that Desgabets was a rationalist. It is one thing to say that the case for empiricism secure, and another to say that the case for rationalism is. Indeed, it isn't all-together clear what rationalism amounts to. An empiricist, for example, can also agree that our thoughts depend on the brain and do not literally "resemble" our sensations, say, in terms of their substantive nature, but perhaps also in terms of the manner in which they bear

[41] Schmaltz (2002, p. 181).

[42] Easton (2002, pp. 205–206).

[43] Lennon and Easton (1992, p. 24).

content.[44] Still it seems plausible that some of Desgabets's metaphysical positions have been mistaken as support for content empiricism. Although Desgabets cautions against falling into obscurity—he confesses to finding Descartes's argumentative style opaque—the presentation of his own position is not a model of clarity.[45] Even so, I do not regard Desgabets as only making metaphysical claims about the causal role of the body. He is taking a stand on the nature and origin of our ideas, and, I contend, there is a definite empiricist flavor present. This isn't to say Cook is completely wrong either, and much of his analysis can be subsumed under an empiricist reading.

So, does he think some form of experience is basic in the formation of ideas? Cook has not shown us the answer is no; I will argue it is otherwise. A more characteristic account of his view is where he says that reasoning "deprived of any *experience*" (my emphasis) *cannot* form ideas of things; this is akin to a canvass "ceasing to be a canvass," representing nothing.[46] I will expand on this theme in two ways: First by showing that Desgabets believes the content of sense experience, not just motions of the body or the sense organs, has a foundational role in the production of ideas and knowledge. Second, I will consider whether he thinks there are no examples of ideas that fail to be derivable from sense experience. Although for the most part I will restrict things to a discussion of knowledge empiricism, it is worth mentioning that Desgabets exhibits several other (arguable) symptoms of empiricism. His emphasis on sense experience as a way of combating skepticism is similar to other early empiricisms. (The mind, especially the understanding, is taken to be a passive faculty until un-interpreted, incorrigible, sense data is received. Even operations of the "pure" understanding are wholly subsumed under internal sensory experience. But the place to start is with his most fundamental principle, his indubitable foundation of all truth.

Desgabets's suggestion for what is known best and most fundamentally is that what he calls a "simple conception," the fundamental operation of the mind is always true[47]—this is offered as a bonafide infallible criterion of knowledge, superior to Descartes's suggestion of clarity and distinctness. Elsewhere, Cook calls this Desgabets's Representation Principle, which again says that a simple conception always has a real and existing object, meaning that it is impossible to think of what does not exist, and it is enough to prove something exists that one can think of it.[48] This is used to show, among other things, why conceiving of

---

[44] E.g., we can all agree that, unfortunately, the word "beer" is not very beer-like. Cf. Easton and Lennon (1992, p. 24).

[45] His discussion of the dependence of all thought on the *body* (*du corps*) in a letter to Malebranche (September 1674) also strikes me as ambiguous in the same way.

[46] S2.IX.iii, my emphasis.

[47] S1Iiv; S2.IV.iii.

[48] Cook (2002).

God, or the external world, for instance, suffices to prove they exist, for one would make fun of a man who would speak about a painting that did not represent anything. However it is no less an absurdity to speak about a thought of nothing, an idea of nothing, a nothing known, etc.[49]

The principle is also highly counter-intuitive,[50] and as noted by Cook, Desgabets displays some ingenuity in defending the position that the Representation Principle is not just true, but indubitably so.[51] My purpose here is mainly to evaluate its significance to content empiricism.

Simple conceptions are also key to understanding the role of sensory experience in Desgabets's epistemology. They are epistemically basic, and serve as foundations for our knowledge about the external world, God, the self, universals, the infinite, and so on. Taken in themselves they cannot lead to error or illusion.[52] Unlike Cook, I contend he also took them to be *perceptual* events—Desgabets sometimes even refers to them as perceptions, as in "our perceptions or simple conceptions … supposes the relation of each perception with its object."[53] This is the decisive point: he thought that they were conscious experiences, either of sensible qualities (e.g., colours, tastes, odors, sounds, etc.[54]), or of "primary" features of the external world (quantity, shape, etc.). Both kinds of simple conceptions are contrasted with "precipitous judgements"—the significance of the latter being reminiscent of Descartes's account of error in the *Meditations*: error results from acts of will that go beyond what is known to be indubitably true. Desgabets seems to share the general structure of Descartes's foundationalism, but he replaces *a priori* clear and distinct ideas with these basic conscious percepts—I am proposing he subscribes to something like a sense data theory where the simple conceptions or perceptions give rise to self-justified beliefs. Allow me to fill in some details.

First by "perceptions" he means sense-experiences. He also calls these "sentiments" or "sensible qualities," the "thoughts and passions" of the soul,[55] these "being nothing but representations of things."[56] Sensible qualities are specifically identified with conscious contents—they refer to the character of subjective

[49] S2.IV.vii.

[50] Desgabets is trying to address a traditional puzzler in the philosophy of language about meaning and fictional things. Desgabets appears to embrace a radical solution sometimes attributed to Anselm and Meinong, namely, that there is a kind of existence possessed by non-existent things (see Lycan 2000, p. 14).

[51] Cook (2002, p. 199).

[52] Cf. Cook (2002, p. 193).

[53] E.g., S1.II.iv.

[54] S1.I.i.

[55] S1.I.ii.

[56] S1.II.iv.

experience, not just the causes of experience; "senses" in contrast, at least sometimes means only "brain processes" causally antecedent to thought:

> [T]he senses excite these perceptions which are mistaken to be corporeal qualities in us … corporeal things … do not have the so-called sensible qualities that are falsely attributed to them. It is thus the soul that knows itself always by the senses, or rather it is man in his whole being, and by consequent it should be said that our perceptions and the soul itself are of sensible things and the proper object of our senses …[57]

He seems to be saying that "perceptions" refers to conscious experience and its representational contents (i.e. sensible qualities). These are caused ("excited") by the "senses" (here "brain processes," though the last occurrence strikes me as ambiguous, insofar as it wouldn't make sense to say that mere causal antecedents have "proper (representational?) objects"). This interpretation is reinforced just slightly further on where he re-asserts that "thoughts and spiritual perceptions" are "excited by the body." There is clearly a distinction being drawn between the motions of the nervous system and the conscious perceptions that result. Of course, so far, one might agree to this while pointing out that this is compatible with even Descartes's views—Desgabets has not yet said that *all* mental activity is a form of sensory perception. However, here and elsewhere, Desgabets contrasts his own view with Descartes's opinion that *perceptions* are epistemically peripheral:

> [F]or after having given admirable lessons to everyone, with respect to the nature of our interior perceptions which according to him are given to us by the senses, he fell again into vulgar thoughts of men who cry against the senses instead of against the precipitation of their judgements and who speak only of pure intelligences disengaged from all commerce with the body.[58]

Desgabets typically contrasts simple conceptions with judgements, especially "precipitous" ones that can lead into error and illusion. But notice in the passage above how it is "interior perceptions" that are contrasted with precipitous judgement—I take it that the former are epistemically reliable while the judgements are not. Yes, the "senses" (physical motions) are in play. However, their specific function is to give rise to these "interior perceptions" or "sentiments" and what is "clear and well known in sentiment is the sentiment or the perception itself, which is all on our side, and which is none other than a thought or an idea, by which we know ourselves intuitively."[59]

A dominant theme in the *Supplement* is Descartes's "great discovery" and "foundation of a true philosophy"[60]—namely sensible qualities, and the recognition that they should not be attributed to material objects. Desgabets immediately

---

[57] S1.I.ii.

[58] S1.II.v.

[59] S1.II.v.

[60] S1.I.i.

spells out six important corollaries of their discovery.[61] These include the location of sensible qualities within the mind alone; that they are sentiments or perceptions, though as I mentioned, he also calls these thoughts; that they do not resemble "the modes or accidents of matter;" that they are nevertheless effects of physical causes; that they cannot be explained within physical theory, though since they wholly reside within the non-physical mind, the material world can be fully explicated through the laws of mechanics and mathematics. This accounts for five of the six implications; however, the remaining one is especially significant. It concerns the nature of knowledge, and the contribution of the internal and external senses. This important consequence of the discovery of sensible qualities begins with Desgabets's reassertion that what is made known by the senses is whatever we are "made [i.e., caused] to think" by physical motion:

> If we are made [by bodily motions] to think of [matter], of rest, of figure, and of all that can result from matter, [the senses] make known to us things that are outside of ourselves, and which are really such as they are known to us by a clear and simple conception, as will be explained later.[62]

Desgabets is saying that this is an *implication* of the discovery that to attribute sensible qualities to objects is to confuse the material and the spiritual. What does he have in mind here? It seems be to that ideas imparted about the "primary" qualities of material objects are always trustworthy, they "are really such as they are known."

If we can avoid making "vulgar judgements" about the function of the senses, then we will discern the actual relationship between sensible qualities and material objects—they are only falsely attributed to them. The "senses"—brain processes—cause "these perceptions," i.e., the sensible qualities we often mistake for qualities of physical objects. It follows that the mind's knowledge of itself is both "by the senses" and, further, "by consequent it should be said that our perceptions and the soul itself are of sensible things and the proper object of our senses."[63] Let it be reminded that this picture of sensible and other qualities[64] is compatible with prototypical empiricists (e.g. Locke), as can be seen by comparing what Desgabets has said with this well-known passage from the *Essay*:

> [T]he ideas of primary qualities of bodies are resemblances of them, and their patterns really do exist in the bodies themselves; but the ideas produced in us by these secondary qualities have no resemblance of them at all. There is nothing like our ideas existing in the bodies themselves.[65]

---

[61] S1.I.ii.

[62] S1.I.ii.

[63] S1.I.ii.

[64] Schmaltz (2002) also suggests a link between Desgabets's realism and empiricism.

[65] Locke (1690), II.viii.15.

Much later in the *Supplement* Desgabets revisits the connection between sensible qualities and acts of simple conception:

> Those who take the trouble to apply this truth [that sensible qualities, including sounds, odors, tastes, warmth, coldness, etc … are "sentiments" wholly in the mind] to subjects arising in the service of life, will have occasion to fight and overturn so many contrary prejudices that it will serve as an admirable exercise to accustom them to suspend the judgement in doubtful things and to distinguish simple conception from precipitated judgements, in what consists uniquely the great rule that must be followed to philosophize well.[66]

The discovery of sensible qualities reveals the deep epistemic significance of conscious experience—sensible qualities guide us in the suspension of precipitous judgement and the formation of simple conceptions, thus leading us to the "great rule" of philosophy: the Representation Principle itself.

Desgabets gives two examples that further illustrate the identification of simple conception with perception: that of imaginary space, and, the feeling of heat. These correspond to outward and inward conscious perception, either of the external world, or, of the soul and its sensible qualities.

In the first, he asks us to consider what occurs mentally when we imagine an unreal space. His answer is that there is simple conception of real space and extension by way of acts of *perception* of the actual environment:

> [I]f we look closely, we see well that the object of their simple conception is space and extension with the dimensions that are perceived there, and that when we speak of space that is nothing, or imaginary space, we form a judgement concerning this space, and we destroy by this what simple conception made us see there, i.e., we form a "being of reason," as we will discuss hereafter.[67]

This helps explain the source of error and illusion: the notion of unreal space comes by way of a further judgement, or act of will, that goes beyond what is given through simple conception.

The second example also attributes error to the will, and clarifies the role of sensible qualities when simple conceptions are formed. When feeling heat from a fire, the simple conception is our conscious awareness of heat. The sensible quality of heat is in the mind, and contrasted with the erroneous judgement that the heat is actually in the fire: "if one says the heat of fire resembles the sentiment that properly is our heat, one adds a judgement to simple perception and falls into the error, because this judgement extends beyond perception."[68] The conception/perception of the felt quality of heat is again something that is not prone to error when it is isolated from further judgements of the will—it is tempting to interpret

[66] S1.I.v.

[67] S2.IV.ii.

[68] S2.I.vii.

Desgabets to be exploiting the assumption that the way things consciously seem is infallible.

Both examples seem to root simple conception in conscious events, implying that external and internal perception are epistemically basic, and reliable, so long as we are careful not to taint them with precipitous judgement: "the confusion of thoughts does not come from our ideas, or from our simple conceptions, but from precipitated judgements which make us say that we see something in our sensations that we don't actually see there."[69] Pursuing a Cartesian theme, he maintains that error is really a misuse of freedom. Deprivation and defect can only be attributed to us, not God, as when we allow our judgements to fail to conform themselves to our raw perceptions. Realizing this, the "alleged deceptions of the senses" can be avoided by

> following the rules that are given for that, of which the principal one is to restrict oneself to that which is known by simple conception, and to suspend one's judgement concerning the remainder, until one has a quite clear idea … If one follows this rule one will not be mistaken even concerning illusions in dreams; and yet, nothing occurs but the true while one stays with the first operation of the mind, which being always true and in conformity with its object as everyone admits …[70]

Simple conception considered in itself can only lead us to reality. And this is his explanation why sense experience considered in itself, that is, separated from the extrapolations and interpretations of the will is a reliable source of knowledge.

Cook doesn't have a lot to say about the nature of simple conception, and he does not offer any reason to think it is not a form of sensory experience.[71] He does agree that the Representation Principle concerns only simple conceptions, not judgements.[72] His discussion also concedes that simple conceptions can at least sometimes take the form of sensations of sensible qualities, such as the feeling of heat.

I agree with Cook that Desgabets draws no distinction between thought *simpliciter* and simple conception; thinking just is conceiving simply, purely, and truly and without judgement. Simple conception is fundamental to understanding, and does not require acts of will. I would add that Desgabets says ideas and thoughts are one and the same as perceptions, for "what there is clear and well known in sentiment is the sentiment or the perception itself which is all on our side, and which is actually none other than a thought or an idea, by which we know ourselves intuitively inasmuch as we are in this state."[73] This idea is wedded to Desgabets's Cartesian division of the soul into an active will and a passive

[69] S2.IX.iv.

[70] S2.IV.viii.

[71] Cook (2002, 2008).

[72] Cook (2002, p. 193).

[73] S1.I.v.

faculty of understanding. The understanding does nothing but receive ideas, sentiments, and knowledge, and these being nothing but "pure passion" in the soul given by the body and the (physical) senses.[74] Even Cook agrees there are no other categories of mentality to worry about: really there is only simple conception, and I am with him when he says that "thought, properly understood just is simple conception."[75] However, in a section titled "that thought, idea, knowledge, perceptions, sentiments, are really the same thing," Desgabets remarks that ideas do not differ in their intrinsic character, but only "accidentally" and "extrinsically" in virtue of their causes.[76] Given that he thinks there are sense experiences, this logically implies simple conception is a form of perception.[77]

Certainly Desgabets's Representation Principle is highly counter-intuitive for at least two reasons. First, it seems we can conceive of things that don't really exist, like a glass of wine that magically refills itself. Second, since conceiving is really just perceiving, again, it seems just false, since the way we perceive the world often deviates from how things really are (as in hallucinations, illusions, and so on). Desgabets repeats the equation of simple conception and perception when he addresses these worries.

First he considers the complaint that the Representation Principle demands something absurd, namely that unreal things, like chimeras, or a physical God, what he calls "beings of reason," aren't really conceived of. But then why does it seem possible to think and speak of them? He responds that "[t]o answer the first difficulty, it is enough to be reminded of what we have said that our principle only applies to simple conception and ideas that we have of things without mixing any judgement which exceeds our perception."[78] So, as with the examples about imaginary space and heat, conceiving simply means not exceeding what is given in sensory perception. If we confine ourselves in this way, thought will always have an existing object—something genuinely perceived (e.g., seen). But the will forms mere "beings of reason," signifying "nothing," when discourses and speech are extended "beyond perception" (or "conception," which he uses as a variant), as in "the discourses that are made by extending speech beyond perception, are not

[74] S2.IV.vi.

[75] Cook (2002, p. 193).

[76] S1.VI.iii.

[77] I acknowledge that this section can be interpreted differently. Perhaps Desgabets is only claiming all mental events are "the same" in that they are all "modes or accidents" of the soul, and that physical motions elicit thoughts by occasional causation. Here he also says that many ideas have "nothing similar to what happens in our brain and in our senses" and "draw nothing at all from any corporeal image." On my reading this only means there is no borrowing or copying of the patterns in physical motions—this is compatible with saying that thought contents are really sensation contents.

[78] S2.IV.viii.

human discourses."[79] Of course, we can *speak* of artificial separations and unions of real things "chimerically," but this is akin to telling a "lie," for him, an assertion of something known or conceived that is not really known or conceived: it is the empty mouthing of words without any content. Setting aside the absurdities of this answer nevertheless leads into another worry. If sense perception is always veridical, how does this square with the many purported examples in which the senses deceive us?

Before continuing, notice that since he is speaking of the *deception* attributed to the senses, the word "senses" undoubtedly takes on an epistemic dimension—it perhaps refers here to sense *experience*, not just physical causes, since it would be bizarre and unintelligible to suppose he means only to resist allegations against deceptive brain processes. Deception, Desgabets says, is not suited to sense perception "at all," for, as Part I of the *Supplement* showed, "it is by their means that we have all the true ideas of things."[80] He then repeats that beings of reason are false judgements extending "beyond perception." Purported examples of deception by the senses are likewise attributed to the non-sensory exercise of precipitous judgement: "the deception that is attributed to the senses is none other than a precipitous judgement by which one says what the sense does not enable one to know."[81] The inferiority of precipitous judgement is contrasted with simple conception, and recall before with "interior perceptions."

Since simple conception is subsumed under the Representation Principle this means Desgabets is also committed to *knowledge* empiricism. This makes sense of Desgabets's claim that "strong proof" there is an Earth, the sun, a God, and so on, is that we "see them, touch them" and so on, and that this is the same as knowing them, which is the same as thinking of them and "form[ing] the idea":

> [I]t is that itself that shows the undeniable truth and the necessity of our [representation] principle that would fall to ground with all that depends on it, if it could happen that the object of our ideas or simple conceptions was not real, i.e., that one could think of nothing. For what other way have we to assure us of the existence of all these things of which I just spoke[82]

Certainty about the external world, God, and so on is guaranteed by the Representation Principle, but its application is directly connected to sensory experience. Sensing appears to be the same as forming self-validating simple conceptions, at least when held apart from precipitous judgement. Simple conceptions, which encompass "all our ideas," are perceptual in nature.

Another of the many puzzling aspects of all this is how we are supposed to know that the Representation Principle is true. It would obviously be self-defeating

[79] S2.IV.viii.

[80] S2.IV.viii.

[81] S2.IV.viii.

[82] S2.IV.v.

(e.g., for a knowledge empiricist) to say that it is known by something like *a priori* pure intellection divorced from sensation. And yet something like this seems to be the only intelligible answer, for it is hard to see what sort of experience could possibly warrant belief in it. Desgabets does say that its denial is an "absurdity" and a "contradiction."[83] He might have in mind the infallibility and incorrigibility of conscious introspection—that it is absurd and contradictory to say that I can be mistaken about the way things seem. However, other possibilities cannot be decisively ruled out. Perhaps he is just insensitive to this difficulty. Or might he think the ultimate foundation of knowledge is *a priori* reasoning after all? Cook's view is that he doesn't so much give an argument for the Representation Principle as to conflate it with intentionality;[84] more charitably, perhaps he means it is an obvious implication of intentionality, properly considered.

These difficulties are not avoided by assuming Desgabets is only a concept empiricist. This is because it is also hard to see what experience could possibly allow him to acquire the idea of the Representation Principle. Perhaps he just means you can't have a sense experience of nothing. Even if an experience fails to represent something in the external world, there are still sensible qualities in experience which represent the soul—this is, I gather, is the job of the "internal senses." He does say at one point, that since the body causes all of our thoughts, this shows that to think of nothing is the same as not thinking at all, and, moreover, this is made known by "experience."[85]

Desgabets's pseudo-justifications remind me of Harman's diagnosis of what he calls the "sense-data fallacy."[86] The fallacy is to assume that even illusory experience must represent something existent—such as sense data—if one assumes seeing always implies a relation to some (possibly mental) object. Harman suggests this fallacious picture comes from focusing on only one side of a linguistic ambiguity: "seeing" can mean either "seeming to see something" or "genuinely seeing something." He calls the later "seeing*". If you think that seeing is always seeing*, then you won't feel at all bad about postulating a zoo of mental entities in order to explain hallucinations and illusions. Perhaps Desgabets is a genuine example in which this linguistic wrinkle is converted into a philosophical mistake, and this is why he appears to believe that no discursive argument is necessary to establish the Representation Principle. This would explain the apparent non-sequitur of supposing that to think of something implies that it exists[87]—the latter signifies what Harman would call thinking*. This analysis is supported by his comparison of thinking to eating. Desgabets argues

[83] S2.IV.vii, S2.VI.vi.

[84] Cook (2002, p. 192).

[85] S1.V.iii.

[86] Harman (1990).

[87] S1.II.vi.

that if "one thought of nothing, one would think without thinking, the same as one would eat without eating if one could eat nothing, and one could say the same of the thought by which the soul is known as a thing that thinks."[88] While the grammar of eating does entail something eaten, this isn't the case for thinking, unless, that is, one follows the assumption that thinking is always thinking*. This is why I say he is a sense-data theorist—though I hesitate to call him an indirect realist. He insists it would be a "bad direction in which this discourse could lead us, if we think that our ideas are like things and that they are intermediate objects between the thought and the object itself of which one thinks." He continues: "[T]he act by which we know an object terminates immediately and directly at the object itself, and not at the idea which is not at all a representative and objective medium, as would be a portrait in which and by the means of which the King would be seen."[89] Combined with the assumption that there is a non-physical soul, and the Representation Principle, this implies that the character of internal experience represents (and resembles) aspects of the soul itself, and, as we'll see below, Desgabets did believe this.

The content-empiricist reading of Degsabets has other virtues, for instance its account of his handling of "pure" intellection. Is there a faculty for knowing that isn't based on sense experience? There isn't: Desgabets thinks that even when the soul knows by pure intellection it be nothing but a sensation that his body gives him. Sensory experience and pure intellection are the *same* thing insofar as sensing directs one to the "true object" of sensory experience, namely the mind and its qualitative character; the soul, by way of its "sensible qualities" is in fact the "true subject of sensation."[90]

It is true, as Cook notices, that pure intellection is not imagistic.[91] For Cook, pure intellection is contrasted with imagination, and these are the only two ways of knowing. In imagination alone Desgabets would say there is an image that bears some resemblance to what is represented. But these remarks are compatible with what I am arguing, since Desgabets thinks that pure intellection is not divorced from the senses either metaphysically or epistemically. As we saw before, ideas or thoughts are really just perceptions or sentiments, and "[i]t is therefore an imaginary thing to make efforts to act without the help of the exterior senses to have pure intellections."[92] When Desgabets says "we have our purest intellections through the senses, the pretended pure understanding distinguished from the senses, is imaginary,"[93] he means there is *both* sense experience and

[88] S1.II.vi.

[89] S1.III.iii.

[90] S1.II.vi.

[91] Cook (2008, p. 514).

[92] S1.I.v.

[93] S1.III.vi, *RD*, p. 186.

causally antecedent brain processes. Pure intellection is a kind of conscious experience—and this fits better with his insistence that at least some experiences, such as of sensible qualities, like the feeling of heat, do not resemble what is represented—as an image would.[94] Desgabets certainly would not accept the claim that all sense experience can be subsumed under imagination. This is why Schmaltz had it right when he said Desgabets recognized no faculty of the understanding that was independent of either sensation or imagination.[95]

Let me briefly restate what has been shown so far. Desgabets's indubitable foundation of all truth is his Representation Principle, which states that all simple conceptions must be true. Simple conceptions are experiential in nature, and even pure intellection turns out to be a form of perception or sensation. Many of these perceptions are of sensible qualities where the soul is "subject and object," meaning the soul is "immediately" and "without reflection" sensing characteristics of itself.[96] Desgabets appears to be a content empiricist who offers an early sense-data theory, motivated by anti-skepticism. These conclusions are compatible with saying that mental processes causally depend on the body.

Actually, it would be more correct to say a firm distinction cannot be drawn between knowledge versus concept empiricism for Desgabets. In light of the Representation Principle, he would say that to form any idea just is to simultaneously guarantee that it is true; there simply are no ideas that fail to be instances of knowledge—ideas that are not knowledge aren't really ideas at all.[97]

The matter as to how we acquire ideas can be further explored, I propose, by following a "Humean" strategy. Although no experience can directly establish that ideas are derivable from something like "simple impressions," one could help establish the plausibility of concept empiricism by showing that even seemingly recalcitrant ideas can be derived from simple impressions after all. Despite Cook's claim that "he does not try to show that apparently problematic ideas like those of

---

[94] S1.III.vi.

[95] Schmaltz (2002, p. 95).

[96] S2.VI.ii.

[97] Others use Desgabets's views about the created eternal truths as grounds for interpreting him as a knowledge empiricist. Ayers (1998, pp. 1029–1030), for instance, argues that he supposed God's voluntary creation of eternal truths implied nothing is necessary, so there is no *a priori* knowledge. Cook rejects this line since Desgabets believes that indeed there are necessary truths (yet somehow "contingent on God"). Easton and Lennon (1992, p. 29) offer another suggestion, which is that since we don't know God's will, we can't know anything outside of experience. Desgabets writes that "to know created things we must wait until God has given them their essence and existence, which are equally contingent." (S2.VIII.vii) (Schmaltz also suggests something like this (2002, p. 16).) However, Cook rejects this claim as well, since it is compatible with this passage that Desgabets only means we use non-sensory intellection to discover the necessary truths that God has created. (I admit to finding this line of argumentation troublesome since Descartes's himself adhered to the doctrine of freely created eternal truths).

being and thought are basic ideas, or can be formed from basic ideas,"[98] the ideas of God, the Soul, and Universals are all discussed in detail. I disagree with Cook's assertion that "Desgabets is saying pure intellections, and thoughts of universal or spiritual things merely depend on physical motions of the body, and not anything more."[99] To show that the content of these ideas are obtained from sensation I won't (and don't need to) rely on ambiguous passages, as when he writes that "the senses are as necessary to have an idea of an angel or a shape in general as to have that of a mountain or a circle."[100] I will argue that, he means sensory experiences, not just generic "brain processes," are necessary.

Consider the idea of God—where does it come from? Desgabets mixes praise with criticism of Descartes's proofs. In the second chapter (Part II) of the *Supplement*, Desgabets concurs with Descartes's conclusion in the *Meditations* that contemplation of the idea of God suffices to demonstrate His existence. However, he complains that the reasoning Descartes offers contains several errors. The "capital defect" is not noticing that the idea of God is not unique in this respect—given the Representation Principle, ideas of things other than God are also self-validating. Does the self-validating nature of the idea of God issue from *a priori* reasoning? No, it does not. It is one thing to say that the idea of God is self-validating. It is another to say it does not depend on sense experience. I have argued that he thinks every idea, including those derived from sense experience, are also self-validating, in that they depend on forming a simple conception/perception.

But what plausible story can be told about how Desgabets thinks it is possible to obtain the idea of God using only sense perception? Would he perhaps agree with Hume that it can be arrived at by amplifying or negating the idea of a limited being, or, might he agree after all with Descartes that the idea is somehow innate? This second option seems to be a non-starter. Cook acknowledges Desgabets's denial that we see "by means of ideas created with us," and speaks of new persons as "blank slates empty of all."[101] Here he would appear to be denying that there are any innate ideas, and obviously this is incompatible with the line of argument Descartes offers in the third meditation. He does say that God "excites in us His idea by a thousand kinds of actions."[102] Admittedly here he might only be referring to the idea's causes. On the other hand, given that motions excite "infinite" perceptions continually experienced,[103] it seems more plausible to read

[98] Cook (2008, p. 506).

[99] Cook (2008, p. 510).

[100] S1.VI.i.

[101] Cook (2008, p. 509, n. 25). See *CdC*, pp. 209–210; *RD*, p. 283.

[102] S2.VIII.x.

[103] S2.XIII.v.

him as saying that (somehow) sense experience, not abstract reasoning, gives rise to the idea (perhaps he means a special kind of God-revealing religious experience).

Continuing, it is by "revelation and tradition," that we obtain ideas about "immaterial" things like angels, and demons, and without which these would be "impossible to think of."[104] He also explicitly connects the acquisition of the idea of God to sensation: "speech which composes the discourses made to us concerning the essence of God and his perfections excite in us the idea of God" just as "the movement of the flame gives us that of heat, depending on what we approach."[105] These too might only be taken as merely a claim about causal antecedents. But when Desgabets also says ideas about God are "infinitely finer and stronger" when caused by "perceptions" such as those found in reading, meditation, and revelation, he would seem to be referring to both conscious experience and its causes:

> [A]ll knowledge acquired by the senses, by reading, meditations, supernatural revelations etc., form a big enough part of the cause of our ideas, for there is no doubt that a man who possesses all of these perfections in a high degree can form thought or ideas infinitely finer and stronger and more understood than another who would not have all these perceptions; there is very true description of this in the Holy Trinity, where the second person who is an infinite and subsistent thought that God forms by knowledge of his own perfections, supposes an infinitely perfect agent, which is the Our Heavenly Father.[106]

It is by "very feeble word … and by instruction" that one obtains "everything that is known of the divine essence, the Holy Trinity, Incarnation and other things that are undoubtedly the greatest objects of our knowledge … simple Christians, even women, conceive all this."[107] Desgabets also addresses, and dismisses, *a priori* approaches to questions of Divinity: "time and trouble would be wasted" in thinking these ideas could be imparted by proof in the style of "arithmetic or geometry." This isn't quite the answer Hume gives, but Desgabets does seem to be saying that what we read, hear, and experience in revelation is necessary for a "more understood" conception of God. Occasionally he is even more forthright. Without sensible qualities:

> we could know neither our soul, nor God, nor angels, and not perceive the connection of our thoughts and movements of body, which appeared so clear to us that all we need to convince ourselves of it is certain reflections upon our interior experience, which produces the clearest possible intuitive knowledge.[108]

This last passage is also a characteristic statement of the soul's knowledge of itself. Here Desgabets is unequivocal about the nature of introspection: "the soul

---

[104] S2.VIII.x.

[105] S1.VI.iii.

[106] S2.IX.ii.

[107] S2.IX.iii.

[108] S1.VIII.i.

knows itself clearly by the senses and perceptions."[109] It knows itself "intuitively as a thing which thinks in an infinity of the means of the senses both internal and external."[110] Section two of chapter six (Part I) is even titled "That the soul is known clearly by senses and perceptions" and again he says these ideas are not just produced by motion, but they are specifically "sentiments and perceptions;" in this way the soul immediately, and infallibly, can attain self-knowledge. This includes awareness of the close union of the soul and body, "as they are experienced continually acting mutually one upon the other,"[111] as well as their "real distinction" as substances.[112]

Certainly these ideas are "excited" by either the exterior or interior senses, nevertheless, as he writes in the *Critique* "knowledge of our thoughts, sentiments and properties of bodies is through experience."[113] The ideas through which the soul knows itself are internally oriented "feelings or perceptions."[114] The soul's knowledge of itself, including its relationship to the body is therefore mediated by "interior experience"[115] or "impressions of the body" such as hunger, heat, and pain[116]—these are the sensible qualities which provide the "clearest possible intuitive knowledge."[117] The soul is the "subject and object" of these feelings, and, as I mentioned before, the "true subject" of sensation.

The soul represents itself through introspective conscious experience; it is "experience that teaches our dependence on the body," for "the faculty that the soul has to know and to want is so vast and that one senses it by an experience so lively and so continual."[118] This leads Desgabets to reject Descartes's claim in the *Meditations* that the cause of an idea must contain as much perfection or reality "formally" as there is perfection or reality "objectively," or representatively, in the idea itself:

> [H]ere is a very general proposition in which I find many difficulties, for I do not understand how it can to be true with regard to ideas that make our soul known to us, because the majority of these ideas are feelings or perceptions which almost never have

[109] S1.II.vi.
[110] S2.IV.viii.
[111] S1.II.ii.
[112] S1.I.iii.
[113] *CdC*., p. 190.
[114] S2.IX.i.
[115] S1.VIII.viii.
[116] S1.VI.ii.
[117] S1.VIII.i.
[118] S1.VIII.viii.

> the external thing that excites them as an object, as we showed above, but which are as many thoughts or of ideas having the soul for an object as having had such a feeling.[119]

The cause of our ideas about the soul are, of course, physical brain processes—but they do not themselves contain the perfections those ideas represent. Sensible qualities represent the soul, not aspects of brain processes:

> However, the efficient cause that created these feelings or perceptions is not the soul at all which is only the subject and object, but rather the external or interior senses aided by the action of surrounding things which produce and excite and form these feelings in us, by a force and very particular property known to us intuitively by experience, so that Descartes's general proposition is not true at all in this respect, for a small movement of a leaf, e.g., cannot produce all the fears that can be found in a surprise, the perils present in the loss of life, nor any others comparable to these great emotions. It must also be said of fire that produces heat, of wine that produces fervor, and other innumerable feelings.[120]

The conscious, and representational, character of intense surprise and fear of hidden danger in no way resembles the physical motions which produced it—the leaf's subtle rustling is not comparable to the "great emotions" which result. Similarly, the sensible qualities excited by fire and wine are not representing something in the fire or in the wine. That is why "[o]ne should also be careful that the ideas excited by the senses are always very clear and very true when one relates them to their own objects, which is often the soul itself in so far as it is in a certain state."[121] In other words, the representational "object"—what the idea is aiming at, is the soul itself.

Desgabets achieved some notoriety for his view that the soul literally instantiates properties represented by sensible qualities. Schmaltz mentions Malebranche's lampooning of this proposal, as if Desgabets were implying it is literally "painted" during experience—"white or black, hot or cold"—and this led to joking about "green souls" and suchlike.[122] But at least this intelligibly reconciles Desgabets's empiricism with the non-resembling character of external sensations. Descartes had used the non-resembling character of sensible qualities to argue that those contents must somehow be already present in the mind—recall the truism that they have to come from somewhere. Desgabets is following a similar path, though I interpret him to be saying, unlike Descartes, that representational contents of sensible qualities are copies of the soul's own properties. This is compatible with Desgabets's denial that our ideas resemble what occurs in the "senses" as "sensible qualities are spiritual perceptions that are on our side and that do not resemble at all what happens in the corporeal senses."[123] A sensible quality does

---

[119] S2.IX.i.

[120] S2.IX.i

[121] S2.XI.i.

[122] Schmaltz (1996, p. 77).

[123] S1.VI.iii.

not copy its physical cause (fire, wine, a leaf moving). But ideas (at least internally directed simple conceptions) do appear to copy or resemble aspects of the soul. This way of accounting for the representational content of ideas might seem radical (even Platonic), but given that he has rejected the possibility of stored contents, what options remain? May it be said that squinting is now required to continue holding the distinction between content empiricism and standard Cartesianism: idea contents are not innate, however, introspection, where the mind scrutinizes itself, derives all the contents it needs from its own intrinsic properties—perhaps we are a short step from calling these stored "contents" rather than "intrinsic properties" immediately available to introspective representation.

Finally, there are ideas of general things. For Desgabets, "the knowledge of universal things is not different from that of particulars save for what is most confused[:]"[124]

> [T]he only difference between the manner in which we use our interior senses to know particular and universal things consists only in that, making the animal spirits upon the corporeal species, we stop them longer upon those [corporeal species] we want to serve to represent something more distinctly and we make them enter there deeply, so that the species recall more strongly and conserve longer the idea that we want to have.[125]

As with others at the time, Desgabets's conception of brain processes is influenced by the example of the circulatory system, and he adopts a hydraulic model in which patterned-tributaries, or "species" spread throughout the brain serve as conduits for the flow of "animal spirits." The patterns themselves are presumably produced by activities of the internal and external senses. However, the will can also exert its spectral influence over the spirits' motions—"the soul acts on the body in voluntary movements, in the same way the body acts on the soul in involuntary thoughts[;]" it can "stop them" and make them "enter deeply" specific pathways.[126] One of the primary functions of the will seems to be the direction of attention. Thoughts endure, for instance, because the will commands the animal spirits to linger at a specific location.[127]

When attention is highly focused, this results in a stronger, more lasting idea with more specificity; however, when the spirits "pass lightly and slide as though superficially" along the hydraulic pathways, we fail to perceive "what distinguishes one from the other," and the idea is more fleeting, less distinct, more confused, and, crucially, more "universal and abstract."[128] This is not exactly to say that universal ideas are derived from ideas about particulars, but they are derived from a similar, though inferior, process.

---

[124] S2.VI.i.

[125] S2.VI.i.

[126] S2.IX.i.

[127] S2.V.ii.

[128] S2.VI.i.

One could attempt to spell this out more concretely, and here is one suggestion. Consider the difference between thinking of a specific glass of wine, and wine in general (i.e., the concept of wine). I imagine a basin within the hydraulic system riddled with fissures. The basins and fissures are the physical manifestation of the relationship between general and particular ideas about wine; basins elsewhere in the system correspond to other ideas, and one can also imagine a labyrinthine, higher-order, and partly overlapping, architecture made up by the various pathways. Perhaps Desgabets is saying that when attention is low, the basin collects only fast-moving spirits, which are distributed widely, and so the fissures are all nearly empty, slight differences between what specific fissures have absorbed are not noticeable, correspondingly, distinctions between particulars are not made. However, since the basin itself contains at least some spirits, these fainter motions still occasion the generic idea of wine—that is, the concept of wine. When the spirits are slower and more abundant, most are deeply absorbed by a specific fissure, and the idea of a particular glass of wine results. This is compatible with saying that a specific idea of the glass of wine presupposes a grasp of the generic concept—a reasonable assumption.

Desgabets's hydraulic theory is severely underdeveloped, and wholly inadequate, not least because it fails to account for creative thought and its infinite capacity to combine and re-combine representations into novel forms. Nevertheless, Desgabets shares the widely held and fundamental interest of his contemporaries in the explanation of human language abilities. Language makes humans highly unusual compared to all other organisms.

The key feature which seems to make it so baffling is that it "permits infinite use of finite means"—this is von Humboldt's expression, a favourite of Chomsky's,[129] which elegantly anticipates what is now described as the productivity and systematicity of thought and language. There are an infinite range of thoughts one can possibly entertain, and sentences one could utter, though, obviously, as a finite organism, a human being cannot actually demonstrate this. Nevertheless, the creative aspect of language use strongly suggests one's finite means can indeed be put to infinite uses—there is an unlimited range of possible representations that one could produce and consume. Only an infinite representational capacity, or competence, can seem to explain how even the most ordinary human being can respond quickly and appropriately to novel linguistic situations (e.g. since they are novel, such feats cannot be explained in terms of a person's prior experience).

Desgabets is also sensitive to this primary concern, as the following passage indicates:

> [T]he simplest among men who have never heard the terms of the School pronounced, have not difficulty at all in conceiving that matter, e.g. could receive different modes and forms successively, and that in fact it always has some one of them, but there is an infinity of others

[129] Chomsky (1995, p. 15).

> such as are enchanted palaces, machines of all sorts of combinations that can be imagined at will …[130]

It important to realize that here he is not just asking how it is possible for the mind to misrepresent (this also continues to be a point of central interest to philosophers). He finds it impossible to see how mere mechanism (governed by exceptionless physical law) could permit the mind to form an infinity of combinations at will, and this motivates the metaphysical separation of the soul from the body. Mechanical principles can provide satisfactory explanations of the behaviour of animals, "all this can be done in beasts without giving them a thinking soul."[131] They are "moved only as automatons or machines," without sentiments and thoughts; similarly "what is done by habit" in humans, likewise, does not demand anything more than an account of physical causes.

For many it remains an orthodoxy that only a compositional system can account for the infinite use of finite means (i.e. one in which syntactic structures are constructed and transformed by formal operations over symbols). Desgabets can be forgiven for not realizing this. Actually, his structured maze of fissures is more reminiscent, though just barely, of connectionism's distributed processing networks, and perhaps anticipates certain aspects of this alternative. The spirits' differing intensities of motion could correspond to degrees of activation in a connection node, or degrees of fit with an activation pattern, perhaps. These comparisons are not to be too seriously pursued, although they do lead one to wonder what he had in mind—a connectionist style architecture tends to be associated with a more empiricist outlook, such as an explanation of language acquisition that does not rely on innate ideas. Perhaps there is also a clue here about the importance of his Representation Principle. Desgabets must have known that its denial can be accommodated by saying that ideas of unreal things are obtained by performing mental operations on representations of real things (e.g. a golden mountain). This is the just the sort of insight that could lead to the development of a compositional theory of representation and thought. Hence, a non-compositional theory must lead in some other direction—perhaps Desgabets realized that if contents aren't to be derived from such operations, then perhaps they can only be obtained if there are always representational objects which literally instantiate the target properties. For this to still count as empiricism, I suppose it would have to be the case that the mind somehow *perceives* both existent and inexistent things. This is weird, but we have seen that weirdness is not a deterrent.

Then again, perhaps here Desgabets is just saying that ideas about universals come from generic brain processes—the maze of fissures is compatible with the rejection of content empiricism. So is his claim that general ideas are inferior, and more confused, than specific ideas. Then again, abstractions are at least one step

[130] S2.VII.ii.

[131] S1.VIII.iv.

removed from sense experience, and so his account of how they are formed is not surprising given content empiricism. Ideas about universals result from low-level activity in the "canals" that might otherwise occasion stronger ideas of particular things. The more you attend to an idea, the more specific it becomes. A decrease in attention corresponds to the occasioning of an idea that is more abstract. This squares with empiricism, if ideas of particulars are obtained by way of sense experience, where attention is under the guidance of the will; this is also probably how the channels are initially formed. There is no special epistemic faculty for acquiring general ideas—they result from a derivative, and inferior, interaction between attention and the movements of the animal spirits through the channels.

As all this talk of wine and spirits is making me thirsty, it is time to bring this discussion to a close.

Conclusion

Desgabets is certainly Cartesian in various ways. However, he does not, like Descartes,[132] dismiss sensible qualities as confused and obscure. He considered Descartes's greatest achievement to be his introduction of sensible qualities to philosophy, but Descartes's overlooked their role in acts of simple conception, the basis of all our ideas, and, indeed, all knowledge. Desgabets thinks we can know what they represent (the soul), whether they represent truly (they do), and why they are important to knowledge (because they make simple conception possible). Sensory experience establishes the existence and nature of body, just as it establishes the existence and nature of the soul, and even Divinity. There is no thought in the absence of sensation.[133]

Cook argues that Desgabets denied ideas must literally resemble, or be copies of, "what happens in the senses."[134] Sensible qualities, like the subjective feel of heat, are in us, not the external world, or even the corporeal sense organs, and so "[k]nowing this, we should no longer assume that if our ideas come from the senses then they must be similar to what happens in the senses … like Descartes, he denies any similarity between our ideas and what happens in the senses."[135] Cook takes this to be incongruent with empiricism; this "is not what one would expect from an empiricist."[136] Actually, this is a non sequitur, and an

[132] AT VII.43; Descartes (1911), 164.

[133] SIIIiv-v, *RD*, pp. 171–174.

[134] Cook (2008, p. 508).

[135] Cook (2008, p. 508).

[136] Cook (2008, p. 506).

ironic one at that, for Cook is now trading on the very ambiguity that he has cautioned against. Cook urges us to interpret Desgabets's use of "senses" to merely mean "motions of the body," as opposed to conscious experience. If this is correct, then all that follows from the non-resemblance of ideas to "the senses" is that they, and indeed, *sensible qualities*—the essential character of experience—do not have to resemble their *physical causes*. Why should that be a reason to deny content empiricism? The non-resemblance of ideas to their physical causes is quite compatible with my thesis, namely that Desgabets thinks the representational content of all ideas—including abstractions, and metaphysical ideas—are exhausted by the contents of sensible qualities.

Cook also addresses the issue of whether an idea must at least resemble its representational object, concluding that Desgabets denies this.[137] However, this is not easy to reconcile with Desgabets's assertion that ideas at least "intentionally" resemble their objects—though admittedly this sort of resemblance is contrasted with "natural," or "real" resemblance, "of being." The distinction here is hard to interpret, and I admit it seems compatible with merely a failure of ideas to resemble their causes. Then again, not once does he claim that ideas fail to resemble aspects of the soul. In the case of fire, recall, the sensible qualities in the idea of heat literally do resemble their object, namely, the soul and its properties. This is why Desgabets's version of the *Nihil* principle is not at odds with empiricism—the change he makes only emphasizes that ideas don't have to resemble their physical causes. Desgabets is saying more than just that ideas depend on motions of the nervous system. The foundational Representation Principle seems to be an empiricist doctrine (in the manner I have defined it), holding that the conscious content of simple conception is obtained either from introspection or exterior perception, separated from interference of the will.

[137] Cook (2008, p. 508, n. 20).

Descartes, Spinoza, and Locke on Extended Thinking Beings

Don Garrett

Introduction

Can we know that nothing is in itself both thinking and spatially extended?[1] This was among the most central and divisive philosophical issues of the early modern period, one with obvious relevance not only to the theoretical understanding of mind and matter, but also to the practical prospects for immortality and, with it, divine sanctions for morality. While many important philosophers—including Nicolas Malebranche, Gottfried Wilhelm Leibniz, and Samuel Clarke—responded affirmatively, perhaps the most famous and influential defense of the affirmative answer was given by René Descartes. While many other important philosophers—including Thomas Hobbes, Henry More, and Anthony Collins—answered negatively, perhaps the two most infamous and influential defenses of the negative answer were given by the two great philosophers born in 1632, Benedict de Spinoza and John Locke.

Descartes' position is expressed clearly in his confident assertion in *Principles of Philosophy* I.53[2] (published in 1644) that thought and extension are "principal attributes" of substances and that every substance has only one principal attribute.[3] The assertion is undefended there, but behind it lie two arguments that he had already presented in the Sixth Meditation of *Meditations on First Philosophy* (1641)—one from separability and one from divisibility. Both Spinoza and Locke studied Descartes' *Meditations* and *Principles* with care; Spinoza even included a version of the argument from separability in his own 1663 axiomatization of

[1] I employ the qualification "in itself" so as to leave aside the question of whether a compound thing can be both thinking and extended in virtue of having a thinking but unextended part and a distinct extended but unthinking part. Descartes, at least, clearly allows that a human being, as a "substantial union" of mind and body, is both thinking and extended in *this* sense. In what follows, I will leave this qualification tacit.

[2] See Descartes (1984/1985/1991, 1996). All subsequent translations of Descartes' texts are taken from the former, which is the standard English edition.

[3] A complication arises from Descartes' doctrine that 'substance' is not applied univocally to God and to created things such as bodies and finite minds, and hence it is not entirely clear whether God has a principal attribute. I will ignore this complication, since Descartes is clear that God is not extended, and his reasons for thinking that God is not extended presumably parallel, at least in part, his reasons for thinking that finite minds are not extended.

J. Miller (ed.), *Topics in Early Modern Philosophy of Mind*, Studies in the History of Philosophy of Mind 9, 85–104.

Descartes' *Principles of Philosophy*.[4] Hence, they must have judged that they could evade the force of these two arguments. Yet neither philosopher directly attempts to diagnose an error in either argument.

I have two primary aims in this paper. The first is to explain precisely *how* Spinoza and Locke, respectively, would have rejected each of Descartes' two famous arguments of the Sixth Meditation. Locke holds that, at least as far as we can tell, created extended thinking substances are entirely possible even if unlikely; but he also argues that no eternal thinking substance is or can be material. Spinoza, in contrast, holds that everything is both thinking and extended, but that no created thing can be a substance. It should not be surprising, then, that their ways of resisting Descartes' arguments differ considerably. Those differences, in turn, motivate my second aim in the paper: to compare and evaluate their strategies for resisting Descartes' arguments against extended thinking beings.

1 The Separability Argument

1.1 The Separability Argument in Descartes

The first and more prominent of Descartes' two arguments about the relation between thought and extension in the Sixth Meditation may be called the "Separability Argument." As written, it is directed at the conclusion that there is a "real distinction" specifically between Descartes' own mind and his own body—that is (as he explains most fully in *Principles of Philosophy* I.60), that his mind and his body are two different substances. His presentation of the argument may be outlined as follows:

> (S1) Everything which I clearly and distinctly understand is capable of being created by God so as to correspond exactly with my understanding of it.
> (S2) [If I can] clearly and distinctly understand one thing apart from another … [then] they are capable of being separated, at least by God. [from (S1)]
> (S3) The question of what kind of power is required to bring about … a separation does not affect the judgement that the two things [that can be separated] are distinct.

[4] The argument occurs as the demonstration of Part 1, Proposition 8 (I p8d) in *Descartes' "Principles of Philosophy"* (*Renati des Cartes Principiorum Philosophiae*), which is included in Spinoza (1985). (All subsequent citations and translations of Spinoza's texts refer to this standard edition.) The presentation corresponds very closely to the specific version that Descartes presents in the axiomatized section that concludes his second set of replies in *Objections and Replies*, published with the *Meditations*.

(S4) [If I can] clearly and distinctly understand one thing apart from another … [then] the two things are distinct. [from (S2) and (S3)]
(S5) [I see that] absolutely nothing else belongs to my [mind's] nature or essence except that I am [i.e., it is] a thinking thing.
(S6) I have a clear and distinct idea of myself [i.e., my mind], insofar as I am [i.e., it is] simply a thinking, non-extended thing. [from (S5)]
(S7) I have a distinct idea of body, insofar as this is simply an extended, non-thinking thing.
(S8) I am [i.e., my mind is] really distinct from my body, and can exist without it. [from (S4), (S6), and (S7)]

Descartes refrains from giving this argument until the Sixth Meditation for three reasons: (i) only in the Third and Fifth Meditations does he argue that an omnipotent God exists, as required for (S1); (ii) only in the Fifth Meditation does he acquire a clear and distinct idea of body, as required for (S7); and (iii) only in the Fifth Meditation does he completely remove the skeptical doubt about whether clear and distinct ideas are true, a doubt that would otherwise call into question the entire argument. Although its stated conclusion and some of its premises are restricted to Descartes' own mind and body, the argument may be readily generalized to conclude that every mind is distinct from every body, simply by replacing his references to his own mind and body with references to all minds and bodies, respectively.[5] Given his view that everything that thinks is thereby a mind and everything that is extended is thereby a body, it follows from the generalized conclusion that there are no extended thinking beings.[6]

1.2 Spinoza and the Separability Argument

Whereas Descartes concludes that every substance has only a single principal attribute—thought for minds, extension for bodies—Spinoza emphatically denies that a substance must be limited to a single such attribute.[7] Thus he writes in *Ethics*:

[5] In saying this, I am assuming that Descartes holds that he can perceive clearly and distinctly the nature of minds generally as well as his own, or at least that he sees that any other mind would be in a position to give the same argument for itself. If he does not hold either of these things, then there is a serious question how he can claim to know that every substance has only one principal attribute that is either thought or extension. I will return to this question in the final section.

[6] Descartes might well be willing to generalize the argument still further to include all possible minds and all possible bodies, so as to conclude that extended thinking beings are not even metaphysically possible.

[7] Spinoza uses the simple term 'attribute' in place of Descartes' 'principal attribute.'

> P10: Each attribute of a substance must be conceived through itself.
>
> Demonstration: For an attribute is what the intellect perceives concerning a substance, as constituting its essence (by D4); so (by D3) it must be conceived through itself, q.e.d.
>
> Scholium: From all these propositions it is evident that although two attributes may be conceived to be really distinct (i.e., one cannot be conceived without the aid of the other), we still cannot infer from that that they constitute two beings, or two different substances. For it is of the nature of a substance that each of its attributes is conceived through itself, since all the attributes have always been in it together, and one could not be produced by another, but each expresses the reality, or being of substance.[8]

Immediately thereafter, Spinoza argues that God, the substance of infinitely many attributes, necessarily exists[9] and is the only substance that exists.[10] He goes on to conclude that thought and extension are among God's attributes.[11] God, therefore, is both thinking and extended.

It is initially surprising that Spinoza grants, in the scholium, that a real distinction between attributes may be conceived, since in Descartes' use of the term a "real distinction" requires two substances and is not a distinction between attributes at all. However, Spinoza's use of the term is readily explicable in light of his parenthetical clarification. Whereas Descartes defines a real distinction as a distinction between two different substances and specifies independent conceivability as a test for a real distinction,[12] Spinoza takes the Cartesian test as constitutive of a real distinction; and since each attribute can (and must) be conceived independently (as *Ethics* I p10 requires), he concludes that there is a (conceived) real distinction between the attributes of a substance despite their being attributes of the same substance.

Where, on Spinoza's view, does Descartes' argument go wrong? Spinoza does not deny (S1) of the Separability Argument; indeed, he holds that whatever can be conceived clearly and distinctly—or, as he more usually prefers to say, conceived "adequately"—actually *has* been created (i.e., caused to be) by God as it is conceived to be, since "God is the efficient cause of everything that can fall under an infinite intellect."[13] Nor would he have any objection to (S3): things that can be separated are not identical, regardless of the power that is required to separate them. Moreover, he agrees with Descartes that we can conceive a thinking substance without employing any conception of extension, and an extended substance without employing any conception of thought; hence, he would not reject versions of (S5)–(S7) generalized to thinking and extended substances, respectively. Rather, the error, for Spinoza, will lie in the inferences that appeal to "clearly and

---

[8] *Ethics* I p10.

[9] *Ibid*. I p11.

[10] *Ibid*. I p14.

[11] *Ibid*. II p1–2.

[12] *Principles of Philosophy* I.60.

[13] *Ethics* I p16c1.

distinctly understanding one thing apart from another." This term, he must say, is ambiguous, for it may refer either to the *separateness of the conceptions* of two things or to the conception of two things *as being separated*. That is, in saying that one can clearly and distinctly conceive of x apart from y, one may mean either:

> (A) It can be that {I clearly and distinctly conceive x} without {I conceive y}.

or

> (B) I can clearly and distinctly conceive {x without y}.

(S2) follows from (S1) only if its antecedent is (B):

> (S2.) If I can clearly and distinctly conceive {x without y}, then God can separate y from x.

(S4), therefore, follows from (S2) and (S3) only if its antecedent is also (B):

> (S4.) If I can clearly and distinctly conceive {x without y}, then x and y are distinct.

Yet (S6) and (S7) make claims not about conceived separation, but only about separate conception—namely, that a mind can be clearly and distinctly conceived without conceiving a body, and a body distinctly conceived without conceiving a mind. Hence, (S8) follows from (S4), (S6), and (S7) only if the antecedent of (S4) is instead understood as (A):

> (S4.) If it can be that {I clearly and distinctly conceive x} without {I conceive y}, then x and y are distinct.

Thus, the argument appears to equivocate on the term 'clearly and distinctly conceive one thing apart from another'. If (A) (i.e., separate conception) does not entail (B) (i.e., conceived separation), then there is no way to get from (S6) and (S7) to the desired conclusion.

Descartes' apparent implicit slide from (A) to (B) may nevertheless seem quite defensible; for if one can clearly and distinctly conceive x without conceiving y *at all*, what possible obstacle could there be to conceiving also that x exists in the *absence* of y? If the clear and distinct conceptions of x and y do not in any way depend on one another, how could x and y nevertheless be so related that even an omnipotent being could not separate them? Indeed, Descartes can be seen as offering just such a response in his *Replies* to the first set of *Objections*, in which Caterus in effect expresses concern about a slide from (A) to (B). By way of dealing with Caterus' example—God's justice and God's mercy, which Caterus claims can be separately conceived without being able to exist apart—Descartes then goes on to explain that his argument requires that the two things in question be separately conceived clearly and distinctly as "complete" beings, rather than as "incomplete" ones, since two beings conceived as merely incomplete may yet prove to depend for their existence on inherence in a substance through which

each must be conceived. Things conceived as substances, he notes—unlike God's justice and God's mercy—meet this conceptual "completeness" condition.[14]

However, this Cartesian defense of the slide from (A) to (B) ignores one crucial alternative: that neither x nor y depends on the other for its existence or conception, and yet that neither one could exist or be conceived to exist in the absence of the other because both are *independently necessary existents* whose non-existence is inherently inconceivable. Since Descartes assumes that all extended substances and all non-divine thinking substances are contingent beings, he silently ignores this alternative. But that is precisely the alternative that Spinoza adopts: since God's thought does not depend on God's extension, nor does God's extension depend on God's thought, either can be readily conceived, for Spinoza, without conceiving the other. Moreover, each conception is "complete" in Descartes' sense, since attributes are conceived through themselves,[15] and not through something else. Yet since the thinking substance and the extended substance both necessarily exist, it is not possible that one should exist *without* the other. They are thus inseparable—and hence, they escape the Separability Argument for their non-identity.

Although Spinoza recognizes only one substance, God, he allows many—indeed, infinitely many—thinking and extended "singular things" (*res singulares*) that are not substances but are instead "modes" of the one substance.[16] Singular things, defined at *Ethics* 2d7 as "things that are finite and have a determinate existence," include, but are not limited to, the human minds and bodies that Descartes intends to include within the scope of his Separability Argument. But although Spinoza grants that only God has a fully clear and distinct idea of any of these singular things as a whole, his explanation of how the Separability Argument goes wrong in application to them would parallel his explanation of how it goes wrong in application to the unique substance. As a thoroughgoing panpsychist, he maintains that things are "animate," though in different degrees.[17] Hence, just as God can be conceived as either a thinking substance or an extended substance, without either conception depending on the other, so too *every* singular thing can be conceived either as a mind or as a body without either conception depending on

[14] Wilson (1978, pp. 191–198) formulates the ambiguity between (A) and (B) and discusses the relevance of Descartes' reply to Caterus at some length. She proposes using the reply to revise the Separability Argument itself fairly substantially. Rozemond (1998, Ch. 1) proposes a very different reconstruction, incorporating the principle that a substance can have only one principal attribute—stated only in *Comments on a Certain Broadsheet* and *Principles of Philosophy*—into the Separability argument itself. Whether these are desirable interpretative reconstructions or not is a question beyond the scope of this paper. I am concerned primarily with the Separability Argument itself, as Spinoza and Locke found it in Meditation Six and as Descartes formalized it at the end of his *Replies* to the second set of *Objections*.

[15] *Ethics* I p10.

[16] They are finite modes that are to some extent "in themselves" and so approximate to being substances in a partial way—*quasi-substances*, as one might say. See Garrett (2002).

[17] *Ethics* II p13s; see also II p7,s and III p1d.

the other. Since, however, there is a necessary parallelism between extended singular things and the ideas—i.e., the minds—of those things,[18] *all* of which follow with equal and absolute necessity from the divine nature,[19] it is not possible for an extended singular thing to exist without the mind of that thing, nor the mind without the extended singular thing.[20] Such a separation is not even clearly and distinctly conceivable, for the only clearly and distinctly *conceivable* ways for thought and extension to be are the (parallel) ways they *actually are*. An extended singular thing and its thinking mind, while separately conceivable, cannot be clearly and distinctly conceived to be separated; hence, they may be—and in fact are[21]—identical.

1.3 Locke and the Separability Argument

Whereas Descartes appeals to God's power to establish that an extended substance *cannot* think, Locke appeals to God's power to establish nearly the opposite: that, at least as far as we can tell, an extended substance *can* think:

> 6. We have the ideas of matter and thinking, but possibly shall never be able to know, whether any mere material being thinks, or no; it being impossible for us, by the contemplation of our own ideas, without revelation, to discover, whether omnipotency has not given to some systems of matter fitly disposed, a power to perceive and think, or else joined and fixed to matter so disposed a thinking immaterial substance: It being, in respect of our notions, not much more remote from our comprehension to conceive, that God can, if he pleases, superadd to matter a faculty of thinking, than that he should superadd to it another substance with a faculty of thinking; since we know not wherein thinking consists, nor to what sort of substances the Almighty has been pleased to give that power, which cannot be in any created being, but merely by the good pleasure and bounty of the Creator. For I see no contradiction in it, that the first eternal thinking Being or omnipotent Spirit should, if he pleased, give to certain systems of created senseless matter, put together as he thinks fit, some degrees of sense, perception, and thought…. What certainty of knowledge can any one have that some perceptions, such as, v.g., pleasure and pain, should not be in some bodies themselves, after a certain manner modified and moved, as well as that they should be in an immaterial substance, upon the motion of the parts of body? … I say not this, that I would any way lessen the belief of the soul's immateriality: I am not here speaking of probability, but knowledge; and I think not only, that it becomes the modesty of philosophy not to pronounce magisterially, where we want that evidence that can produce knowledge; but also, that it is of use to us to discern how far our

[18] *Ibid*. II p7.

[19] *Ibid*. I p33.

[20] Since singular things come into existence and go out of existence, it is important to distinguish the question of whether their existence *at the particular times at which they exist* is necessary or contingent, from the question of whether there is any time such that it is possible for them *not to exist at that time*. Singular things lack necessary *eternal* existence, for Spinoza, but they do not lack a necessary *durational* existence derived from the necessity of their causes.

[21] *Ethics* II p7s.

> knowledge does reach; for the state we are at present in, not being that of vision, we must, in many things, content ourselves with faith and probability; and in the present question, about the immateriality of the soul, if our faculties cannot arrive at demonstrative certainty, we need not think it strange.[22]

Locke, like Descartes and Spinoza, characterizes some ideas as "clear and distinct," but he understands the distinctness of ideas rather differently, primarily in terms of the fixedness of their relation to terms signifying them.[23] It is not clear that Locke would grant (S1) as Descartes formulates it, since conception might be clear and distinct in Locke's sense and yet sufficiently partial as to hide a contradiction or impossibility. More to our purpose, however, Locke also has a notion of "adequacy" for ideas, which he explains as the perfection of an idea's representation of its archetype.[24] Let us suppose, therefore, that he interprets "clear and distinct understanding" throughout the Separability Argument as "understanding using adequate ideas." Since he characterizes God as omnipotent (for example, at *Essay* IV.x.13), it seems likely that he would grant it to be in God's power to create whatever can be *adequately* conceived, at least; hence, he would not object to this version of (S1). In addition, he would presumably allow that whatever things can be separated by any power are distinct from one another, and so would not object to a parallel version of (S2). Perhaps he would object, as Spinoza must, to the apparent equivocation involved in the inferences from (S2) to (S4) to (S8). But as the cited passage indicates, Locke's central objection, unlike Spinoza's, would surely be to the introspective claims made in (S5)–(S7).

Locke does not, of course, deny that minds, his own included, are things that think. They are, as he sometimes puts it, "cogitative" beings. He does deny, against Descartes, that cogitative beings must *always* think; it is, he claims, no more necessary that a cogitative being always think than that an extended being always move. Hence, constant thinking, at least, cannot be essential to such a being. But even assuming that 'thinking thing' means merely "a thing that can think," Locke would still object to (S5)–(S7). In order to understand that objection, it is necessary to understand something of his conceptions of *substances* and *essences*.

According to Locke, we conceive of substances, of whatever kind, by combining the "obscure" and "relational" idea of "substance-in-general" with ideas of particular qualities. This idea of substance-in-general is the idea of a support of qualities, "some substratum wherein they do subsist, and from which they do result."[25] The idea of "body" or "material substance," for example, results from combining the idea of substance-in-general with the idea of extension and the idea of "solidity"[26]—solidity being the quality whereby bodies exclude other

---

[22] Locke (1975) (*Essay* IV.iii.6). All subsequent citations of Locke's texts refer to the standard edition.

[23] *Essay* II.xxix, "Of Clear and Obscure, Distinct and Confused Ideas".

[24] *Essay* II.xxxi, "Of Adequate and Inadequate Ideas".

[25] *Essay* II.xxiii.1.

[26] *Essay* II.xxiii.15, 27.

bodies from the places they occupy.[27] But there is nothing in any of these ideas that prevents their combination with an idea of thinking into a single idea of a substance or that renders those ideas (in Locke's phrase) "repugnant" to one another. So far, then, we can see no reason why the creation or generation of a thinking material substance should be outside the reach of God's omnipotence. While we cannot see specifically *how* thinking and extension could be combined in a substance, this is not surprising given the obscurity of our idea of substance-in-general and our ignorance of the way in which the qualities of minds and bodies, respectively, "result" from the substrata in which they "subsist." We are, as Locke remarks in the quoted passage, equally unable to see *how* motions of bodies could produce sensation in an *unextended* substance; yet that must happen somehow if our sensing minds are not extended. For all we know, then, our minds may be extended thinking substances.

Locke goes on to distinguish two kinds of essences, "nominal" and "real."[28] A *nominal* essence is that which makes a thing be classified as belonging to the sort or species that it does—thus, he asserts, a nominal essence is an abstract idea, often combining ideas of several qualities, and signified by a general term. A *real* essence, in contrast, is the real internal constitution of a thing from which its "properties" (i.e., "*propria*," a technical term designating constant qualities following unchangeably from an essence) "flow." Hence, thinking belongs to the nominal essence of "cogitative beings" considered as such (i.e., under that abstract idea); and, indeed, nothing else belongs to that particular nominal essence. In this sense, and thinking of one's mind simply as a "thinking" or "cogitative" being, (S5) is true: one may well see that nothing belongs to "the mind's" nominal essence other than thinking. But this is simply an arbitrary classificatory point; any particular cogitative being also falls under many *other* kinds, each with its own abstract idea serving as its nominal essence. From this nominal-essence version of (S5), an acceptable version of (S6) would not follow, for it does nothing to show that an *adequate* idea of any particular thinking substance (including one's own mind) would represent that substance as *unextended*—the idea of extension being fully compatible with the ideas of thinking and substance-in-general. If, on the other hand, we interpret (S5) as a claim about the real essence of particular thinking substances such as one's own mind, then (S5) will simply be false. For Locke claims that we cannot determine whether or not a particular finite thinking substance is a material substance to which God has "superadded" the power of thinking; and if it is such a material substance, it already has the nature or essence of a material, and hence extended, substance as well. (Indeed, it is not immediately clear whether thinking, or the power of thinking, would become even a part of its real essence, as opposed to being an accidental and transitory quality.)

---

[27] *Essay* II.iv, "Of Solidity".

[28] *Essay* III.iii.15–17.

Locke would also object to (S7) on similar grounds. For although one can, without contradiction, form an idea of a body—i.e., an extended, solid substance—without conjoining the idea of thinking to it, there is no guarantee that such an idea will be a distinct or adequate idea of any particular body. On the contrary, if God has superadded the ability to think to a body, then an adequate idea of that body, at least, will have to include an idea of that power.

2 The Divisibility Argument

2.1 The Divisibility Argument in Descartes

The second of Descartes' two arguments concerning the relation between thought and extension in the Sixth Meditation may be called the "Divisibility Argument." It occurs in the course of his explanation of sensory error. His confidence in it, however, is indicated by his remark that "this one argument would be enough to show me that the mind is completely different from the body, even if I did not already know as much from other considerations." The argument, as he presents it, may be outlined as follows:

(D1) If a foot or arm or any other part of the body is cut off, nothing has thereby been taken away from the mind.
(D2) It is one and the same mind that wills, and understands, and has sensory perceptions.
(D3) The faculties of willing, of understanding, of sensory perception, and so on … cannot be termed "parts of the mind." [from (D2)]
(D4) When I consider the mind, or myself insofar as I am merely a thinking thing, I am unable to distinguish any parts within myself; I understand myself to be something quite simple and complete. [from (D1)&(D3)]
(D5) The mind is utterly indivisible. [from (D4)]
(D6) There is no corporeal or extended thing that I can think of which in my thought I cannot easily divide into parts.
(D7) The body is by its very nature always divisible. [from (D6)]
(D8) The mind is completely different from the body. [from (D5) and (D7)]

As with the Separability Argument, the conclusion may be generalized to all minds and bodies;[29] and given the Cartesian doctrine that everything that thinks is

[29] As with the Separability Argument, Descartes may also be willing to generalize the argument further, to all *possible* minds and all possible bodies.

thereby a mind and everything that is extended is thereby a body, it follows that there are no extended thinking beings.

2.2 Spinoza and the Divisibility Argument

As we have seen, Spinoza's rebuttal of the Separability Argument takes a single general form whether its scope is taken to be substances or singular things: in each case, the extended thing and the corresponding thinking thing, while independently conceivable because involving different attributes, can neither exist apart nor be conceived to exist apart because each exists necessarily whenever the other does. In contrast, Spinoza's strategy for rebutting the Divisibility Argument will differ depending on whether it is taken as an argument concerning substances or singular things. This is because he regards substance as indivisible, but at least many singular things—namely, those he also characterizes as "individuals" (*individua*)[30]—as divisible.

Spinoza argues for the indivisibility of substance in *Ethics* I p12 and I p13. The first of these propositions denies that a substance can be divided into its attributes, while the second denies that a substance can be divided within any of its attributes:

> P12: No attribute of a substance can be truly conceived from which it follows that the substance can be divided.
>
> Demonstration: For the parts into which a substance so conceived would be divided either will retain the nature of the substance or will not. If the first [NS[31]: viz. they retain the nature of the substance], then (by P8) each part will have to be infinite, and (by P7) its own cause, and (by P5) each part will have to consist of a different attribute. And so many substances will be able to be formed from one, which is absurd (by P6). Furthermore, the parts (by P2) would have nothing in common with their whole, and the whole (by D4 and P10) could both be and be conceived without its parts, which is absurd, as no one will be able to doubt.
>
> But if the second is asserted, viz. that the parts will not retain the nature of substance, then since the whole substance would be divided into equal parts, it would lose the nature of substance, and would cease to be, which (by P7) is absurd.
>
> P13: A substance which is absolutely infinite is indivisible.
>
> Demonstration: For if it were divisible, the parts into which it would be divided will either retain the nature of an absolutely infinite substance or they will not. If the first, then there will be a number of substances of the same nature, which (by

[30] In *Ethics* II p13s, Spinoza mentions an "infinite individual" composed of all finite individuals. *This* individual would not be a "singular thing," since singular things are by definition finite.

[31] 'NS' indicates an interpolation from the *Nagelate Schriften*, the Dutch translation of Spinoza's *Opera Postuma* prepared by his friends from his Latin manuscripts.

> P5) is absurd. But if the second is asserted, then (as above [NS: P12]), an absolutely infinite substance will be able to cease to be, which (by P11) is also absurd.
>
> Corollary: From these [propositions] it follows that no substance, and consequently no corporeal substance, insofar as it is a substance, is divisible.
>
> Scholium: That substance is indivisible, is understood more simply merely from this, that the nature of substance cannot be conceived unless as infinite, and that by a part of substance nothing can be understood except a finite substance, which (by P8) implies a plain contradiction.

It is clear from these arguments how Spinoza would object to a version of the Divisibility Argument formulated as an argument specifically about substances. While granting (D5), that a thinking substance is utterly indivisible, he would deny (D6), and hence also (D7), by insisting that there *is* an extended substance—indeed, the *only* extended substance—that he cannot conceive to be divided. For although various operations might be properly conceived as dividing a singular thing into parts, no such operation would introduce any division into infinite extended substance itself. An extended substance, as extended, must have *regions*, of course (or, better, *be regionalized*), but these regions are not *parts*, in the sense of things prior to a whole out of which they are generated by composition (nor, indeed are they *things* at all by Spinoza's standards); and any alteration of the modes of the substance is merely a qualitative regional change, not a division into parts.

As noted, those singular things that are composed of parts are *individuals*, in Spinoza's terminology, and these include human beings.[32] To a version of the Divisibility Argument formulated in terms of individuals rather than substances, Spinoza would respond by granting (D6) and (D7) while denying (D5), the claim that minds are utterly indivisible. For the parallelism of thought and extension, according to which "the order and connection of ideas is the same as the order and connection of things"[33] entails that the human mind—which is the idea of the human body—is literally composed of ideas of the parts of the human body. This is clearly stated in *Ethics* II p15:

> P15: The idea that constitutes the formal being [*esse*] of the human Mind is not simple, but composed of a great many ideas.
>
> Demonstration: The idea that constitutes the formal being of the human Mind is the idea of a body (by P13), which (by Post. 1) is composed of a great many highly composite Individuals. But of each Individual composing the body, there is necessarily (by P8C) an idea in God. Therefore (by P7), the idea of the human Body is composed of these many ideas of the parts composing the Body, q.e.d.

[32] *Ethics* II p13s. For purposes of citation, I am treating the so-called "Physical Digression" that precedes *Ethics* II p14 as part of II p13s.

[33] *Ethics* II p7.

Rejecting (D5), of course, requires rejecting (D4) as well—and Spinoza emphatically does so. This does not mean that he denies (D2) or (D3): willing and understanding are not at all distinct parts of the mind for him, and while he writes as though intellect and imagination (which includes sensory perception) can be considered as "parts" of the mind, he does not suppose that one could have imagination without any intellect at all.[34] But all human thinking is awareness of one's own body, on Spinoza's view, and one's various ideas of how things are in the various parts of one's body do constitute parts of one's mind. Hence, he would deny (D1): the removal of a body part would necessarily be paralleled by the removal of the part of the mind that is the idea of that body; and an idea of that body part, although perhaps no longer part of a finite mind having as much consciousness as the human mind, would continue to exist as a singular thing and as a mode of thinking of the one substance.

Because all singular things have minds, for Spinoza, similar points apply to all non-human individuals as well. In addition to discussing the complex singular things that are individuals, however, he also writes in *Ethics* II p13s of the "simplest bodies" (*corpora simplicissima*) that are their ultimate constituents, distinguished from one another *only* by motion-and-rest. These simplest bodies presumably satisfy the definition of 'singular thing' at 2d7: they are finite, clearly have a spatially determinate existence, and are not said to be everlasting. Like other singular things that are modes of extension, then, they too must have corresponding ideas that are their minds. Simplest bodies, as modes of extension, are not unextended—indeed, they may well have various shapes and sizes—but they are spatially homogeneous distributions of different degrees of "motion-and-rest" (*motus & quietis*), the fundamental pervasive feature of infinite extended substance by which that substance is variegated.[35] Spinoza does not explicitly state whether he regards simplest bodies as divisible or not. If he does regard them as divisible (if, for example, they can be split into two smaller simplest bodies by collision), then his response to a version of the Divisibility Argument directed at them and their minds will parallel his response to the Divisibility Argument directed at individuals: both the simplest body and its mind will be equally divisible. If he does not regard them as divisible, then his response will parallel his response to a version of the Divisibility Argument directed at substances: neither the simplest body nor its mind will be divisible.[36] In either case, the Divisibility Argument is blocked.

---

[34] *Ethics* II p46.

[35] For a fuller account, see Garrett (1994).

[36] Presumably this would be Spinoza's response for the "infinite individual" (composed of all other individuals) mentioned in *Ethics* II p7, since he is unlikely to regard it as divisible, despite its composition of parts.

2.3 Locke and the Divisibility Argument

Much as in the case of Spinoza, it will be useful to distinguish Locke's response to the Divisibility Argument as it applies to created things from his response to it as it applies to an eternal substance. Let us consider first the application to created things.

Locke must allow (D2) and (D3), for he emphasizes just as much as Descartes does that the various faculties of the mind are not distinct "agents" or "real beings"—they are mere powers or capacities of one thinking agent that has a variety of ideas and volitions.[37] Locke also appears not to dispute (D7), writing, for example, that "in any bulk of Matter, our Thoughts can never arrive at the utmost *Divisibility*, therefore there is an apparent Infinity to us also in that …."[38] Locke's objection to the Divisibility Argument in the case of created beings—like Spinoza's in the case of created finite individuals—must therefore be to (D5) and, with it, to (D4) and (D1). Unlike Spinoza, he does not claim to be able to *discern* parts in the created mind;[39] but he will not allow that it follows from this that what thinks in him is definitely *not* a divisible system of bodies. For he claims no introspective or other access to the nature of the substance that thinks in him, beyond knowing that it sustains and supports his thoughts and volitions. This substance may be a brain or a "System of fleeting animal spirits[;]"[40] and he even considers, in his discussion of personal identity, the possibility that a separated "little finger" might retain some consciousness.[41] Hence, we cannot know that all thinking beings are indivisible.

While Locke expresses openness to the possibility of created material thinking beings, however, he devotes considerable attention, at the conclusion of his demonstration of the existence of God in *Essay* IV.x ("Of the Existence of a GOD"), to arguing that there is no eternal material being—and especially not an eternal material *thinking* being.[42] Having demonstrated to his own satisfaction that there is an "eternal "most powerful and most knowing" being—namely, God—he reiterates that, just as "nothing" cannot give rise to "something" and what is "purely matter" cannot possibly give rise to motion, so "bare matter" cannot, even if it is in motion, of itself give rise to thought or any thinking thing. By 'purely matter,' he evidently means *having just the basic material qualities of extension and solidity, plus whatever these necessarily entail*; and by 'bare matter', he appears to mean *pure matter with or without motion added to it*. It is "impossible to conceive" that bare matter could ever "have originally in and from it self Sense,

[37] *Essay* II.xxi.6.

[38] *Essay* II.xvii.12.

[39] It is worth noting, however, that he does regard all body parts as parts of himself as a *person*, even if they are not parts of his mind (*Essay* II.xxvii.17–21).

[40] *Essay* II.xxvii.17–21.

[41] *Essay* II.xxvii.13.

[42] For useful discussion of this argument, see Wilson (1979, 1982) and Ayers (1981).

Perception, and Knowledge," he argues, for if it could do so then "Sense, Perception, and Knowledge must be a property [in the technical sense noted earlier] eternally inseparable from Matter and every Particle of it."[43] To see that these qualities are *not* distributed to every particle of matter, he claims, we need only note that, despite the common tendency to think of "matter" as a single thing, it is in fact an infinite number of material particles, so that to allow bare matter to be an eternal thinking thing would require an infinite collection of limited thinkers that would be "independent one of another, of limited force, and distinct thoughts" and hence could not be the source of the "order, harmony, and beauty" that we find in Nature.[44] After noting that "whatsoever is first of all Things, must necessarily contain in it, and actually have, at least, all the Perfections that can ever after exist"—including thought—he concludes that "the first eternal Being cannot be Matter."[45] By this he means, presumably, that it cannot be "pure" or "bare' matter.

Locke then considers two alternative hypotheses according to which something material would nevertheless be eternal. The second of these is not directly relevant to our main question; it is the hypothesis that matter, even if non-thinking, might still be eternal *in addition to* a separate eternal but immaterial thinking being.[46] The first hypothesis, however, is highly relevant: that an eternal thinking being—which Locke assumes would be God—might, even if not deriving its thought just from its purely material nature, nevertheless *have* a material as well as a cogitative nature.[47] Even if not pure matter, it would nonetheless be some kind of thinking matter.

In order to refute this hypothesis, Locke divides it into three alternatives: (i) that *all* matter is eternal and thinking; (ii) that *one single atom* of matter is eternal and thinking; and (iii) that some particular *system* of material particles is eternal and thinking even though its individual elements do not think.[48] Against the first alternative, Locke claims that the result would be an "infinity of Gods," something which defenders of eternal thinking matter will "scarce say." Against the second alternative, which he declares to have "as many Absurdities as" the first, he offers

---

[43] *Essay* IV.x.10.

[44] Spinoza would not insist on the "order, harmony, and beauty" of Nature, since he sees these characteristics merely as projections of human sensibility (as explained in the Appendix to Part 1 of the *Ethics*). He would also allow that thought is not the *consequence* of an extended nature. However, he would insist that every particle of matter is a mode that necessarily also thinks; and while these modes are indeed "limited," they are not "independent," since they are modes (not parts) of one infinite and eternal thing having the utmost perfection and reality.

[45] *Essay* IV.x.10.

[46] *Essay* IV.x.18–19.

[47] *Essay* IV.13–17.

[48] Notably absent from this list is the alternative that more than one eternal atom of matter thinks while other atoms do not. Presumably, however, Locke would make basically the same objections to this alternative that he makes to the alternative that only one atom thinks: either the thinking atoms are the only eternal ones, in which case they create the others, or the thinking eternal atoms differ from the unthinking ones for no reason.

a dilemma: either this single thinking atom is the only eternal thing or it is not. If it is the only eternal thing, then it must create all other matter—doing so, presumably, by its powerful thought, since this will be its only evident difference from other matter. Accordingly, the friends of eternal thinking matter will be under pressure to admit, against their inclination, that some matter has been created by thought, and they will in any case be forced to give up their "great Maxim" that *ex nihilo*, *nihil fit*. Yet to maintain that the single atom is not the only eternal thing would be to hypothesize "without any the least appearance of Reason" that this one atom vastly surpasses the other eternal things. Finally, against the third alternative, Locke has two objections. First, it makes wisdom dependent on the mere juxtaposition of parts, whereas in fact it is "absurd" that any mere position of parts of matter could ever produce thought and knowledge. Second, the parts of such a system must either be at rest or in motion; but if they are at rest, the system is a mere lump equivalent in power to a single atom, while even if they are in motion, wisdom still cannot arise from the "unregulated" and "unguided" motions of the individual parts.

Since these arguments appeal prominently to the thesis that all material things have material parts, it may appear that Locke is offering his own restricted analogue of the Divisibility Argument: a version intended to demonstrate that, while all matter is inherently divisible, this divisibility in an eternal being is incompatible with thinking, so that any eternal thinking being must be unextended. The appearance is heightened by his references to the "impossibility of conceiving" bare matter to have thought "from itself" and to the "absurdity" of the three alternative versions of the more general hypothesis that some eternal matter thinks.

This appearance is deceptive, however, for several reasons. First, Locke is best understood as arguing only that there *is* no eternal material cogitative being, not that such a being is literally *impossible*. For example, it is a key premise of the argument against thinking bare matter that an infinite number of finite Gods could not produce the order, harmony, and beauty that we actually see in nature, and this is presumably also the source of the "absurdity" of the first version of the more general hypothesis of eternal thinking matter. But in the absence of a further argument that such order, harmony, and beauty are themselves necessary and not merely contingent features of the universe, any argument relying essentially on this premise can at most show that an infinity of Gods is not actual. Since Locke explicitly declines to endorse the ontological argument for God's existence,[49] such a further argument does not appear to be forthcoming. Furthermore, the fact that a single eternal thinking atom would require the friends of eternal thinking matter to "allow" the creation of matter by thought and give up "their favourite maxim" is purely *ad hominem*; and the apparent absence of a reason *why* only some eternal atom (or atoms) among others should think does not show the *impossibility* of

[49] *Essay* IV.x.7.

such an atom (or atoms) on any stated Lockean principle.[50] If there is no internal contradiction in the supposition that a cogitative and a material nature are combined in a single substance, then it is hard to see how there could be a contradiction in the supposition that they have eternally been so combined. While Locke might well have *wanted* to be able to argue that eternal thinking matter is impossible, he simply lacks the resources to do so.

Indeed—and this is a separate point—it is not clear that Locke is really even claiming to have *knowledge*, in his strict sense of the term, as opposed to probable opinion,[51] that there is no eternal material thinking thing. For despite his frequent invocations of "absurdities" in his opponents' position, his response to the objection that from God's existence "it does not follow, but that thinking Being may also be material" begins, "Let it be so …."[52] Furthermore, he begins his three-part discussion of the general hypothesis of eternal thinking matter with the mild proposal: "But now let us see how they can satisfy themselves, or others, that this *eternal thinking being is material*."[53]

More important, however, and perhaps more surprisingly, Locke does not deny that God, the eternal thinking being, is extended. As we have seen, materiality—i.e., being a body—requires both extension and solidity, according to Locke. He certainly denies that God has solidity; unlike Descartes, however, he allows that things can be extended without being bodies.[54] This is perhaps most evident in his treatment of space, which he allows to be extended without being a body. But it is equally true of his account of God's location, as he presents it in *Essay* II.xxvii, "Of Identity and Diversity." For all identity requires, on his account, "Existence it self, which determines a Being of any sort to a particular time *and place* incommunicable to two Beings of the same kind."[55] Whereas immaterial finite spirits are located without being extended,[56] God is "without

[50] Locke's own causal maxim, that "a cause is required for every beginning of existence," could not establish such an impossibility, since we are concerned with an *eternal* cogitative atom.

[51] Knowledge is "the perception of the connexion and agreement, or disagreement and repugnancy, of any of our ideas" (*Essay* IV.i.1) and is limited to intuition, demonstration, and sensation. Probability is "the appearance of such an agreement or disagreement, by the intervention of proofs, whose connexion is not constant and immutable, or at least is not perceived to be so, but is, or appears for the most part to be so, and is enough to induce the mind to judge the proposition to be true or false, rather than the contrary." (*Essay* IV.xv.1) If Locke's arguments are meant to provide probability rather than knowledge, they must exemplify one or both of his two "grounds of probability": conformity to past experience and testimony.

[52] *Essay* IV.x.13.

[53] Italics in original.

[54] *Essay* II.xiii.16: "Who told them, that there was, or could be nothing, but solid Beings, which could not think; and thinking Beings that were not extended? Which is all they mean by the terms *Body* and *Spirit*."

[55] *Essay* II.xxvii.3, emphasis added.

[56] *Essay* IV.iii.6.

beginning, eternal, unalterable, and *every where*."[57] Locke's attribution to God of literal omnipresence—and not merely a figurative omnipresence through the effects of divine power—is confirmed by his pointed recommendation in *Essay* II.xiii.26 that we consider very seriously whether the words of "the inspired philosopher St. Paul" that it is "in God" that "we live, move, and have our being" should not be understood literally. God is thus co-located with bodies and also with immaterial finite spirits; but as the passages already cited from *Essay* II.xxvii.1–3 indicate, Locke has no objection to co-location of substances, as long as the substances are not "of the same kind." Locke does propose at one point that we may, if we wish, limit the term 'extension' to bodies, adopting the term 'expansion' for other spatial things; but he admits that, whichever term we use, we are signifying the same idea.[58]

Thus, while Locke denies that there are in fact any eternal material thinking beings, he can and should resist a version of the Divisibility Argument restricted to eternal things. First, it is not clear that he would claim to know even analogues of (D4) and (D5) that were restricted to eternal thinking things. Second, he would reject (D6)'s casual identification of extension with corporeality (i.e., materiality), and he could easily maintain, against (D6), that God is, in light of His omnipresence, both extended and indivisible. Most importantly, however, he would also resist the inference from an analogue of (D8), asserting that no eternal thinking substance is a body, to the conclusion that there cannot be an eternal extended thinking being; for he rejects the principle that every extended being is a body.

3 Evaluation and Conclusion

Many philosophers have found Descartes' doctrine that there are no extended thinking things deeply attractive. Others, including most contemporary philosophers, have found it to be objectionably anti-naturalistic. We may distinguish two broad strategies for denying that it can be established. The first strategy involves defending, primarily, an account of our cognitive faculties from which it follows that the doctrine cannot be known to be true. The second strategy involves defending, primarily, a broader positive metaphysics according to which the doctrine is definitely false.

Locke's approach to the issue is a prime example of the first strategy. His accounts of our limited conceptions of substance and essence leave us without the resources to establish the truth of Descartes' key premises about thinking and material substances in both the Separability Argument and the Divisibility Argument. The central advantage of Locke's approach is that it puts the burden of proof on

[57] *Essay* II.xxvii.2; emphasis added.

[58] *Essay* II.xiii.27.

Descartes to explain how we can have the kind of knowledge required to support his conclusions.

At the same time, however, Locke's position is also subject to serious limitations. He grants that the kinds of powers bestowed by a material nature seem to us naturally inadequate for thought. Moreover, he seems to concede—especially in his objection that eternal thinking pure matter would require an infinity of "independent" thinkers of "distinct" thoughts—that we cannot comprehend how something divisible could be a unified subject or bearer of thoughts.[59] For while he proposes that God might be able to bestow the power of thought on a system of material bodies, he does not explain how God would bring it about that thought was a quality or modification of that *entire system* of bodies and no other. At least as we conceive things, for Locke, it is not sufficient simply for God to create thought; God must provide for some substratum in which that thought subsists. We can readily understand predications of qualities to complex material things—say, a shape or motion to a tree—in virtue of the qualities of the whole resulting simply from the combined qualities of the individual parts, parts themselves considered as substrata. But since Locke rejects panpsychism, it seems that he cannot avail himself of this strategy in the case of complex thinking things. The only alternative seems then that God (or perhaps just an eternal arrangement of things) must have specially constituted a particular system of bodies as a basic substratum *in its own right*, giving it the kind of unity that is evidently required for a mind. The deficiency of our idea of substance-in-general, however, prevents us from seeing how, or even whether, this can be so. Locke's ultimate reply to objections to the effect that it is difficult to see how the materialist scenarios he considers could be realized is simply that it is also difficult to see *how* the *alternatives* to those scenarios could be realized either.

Spinoza's approach, in contrast, is an example of the second strategy. Whereas Locke's overall position is subject to criticisms derived from the modesty of his epistemic resources, Spinoza's is subject to criticisms derived from the strength of his metaphysical claims. He rejects the Separability Argument by holding that there is necessarily a substance with multiple separately conceivable attributes, including thought and extension. To Descartes' predictable objection that it is impossible for one thing to have two different "natures,"[60] he will reply that the perfection of the necessarily existing divine substance, as established by the ontological argument and the principle of sufficient reason, actually requires that one substance have all *possible principal* attributes, necessarily mirroring one another. Since Spinoza agrees with Descartes that thought and extension *are* principal attributes, this means that the one substance and each of the singular

[59] For a compelling contemporary presentation of a related problem about how thoughts could belong to a concatenation of physical particles, see Unger (2006, Ch. 7). He calls this problem the "Experiential Problem of the Many."

[60] Descartes makes this claim in *Comments on a Certain Broadsheet*. See Rozemond (1998, Ch. 1) for discussion.

things that are its finite modes must exist in multiple fundamentally different yet complementary dimensions of being—including as thinking and as extended. This is nothing less than panpsychism, a strong and counterintuitive claim indeed.

Spinoza's response to the Divisibility Argument equally implies panpsychism, for it depends on his doctrine that every individual thing with extended parts has a "mind" whose thinking parts are the minds of those parts. His response also implies that the very same idea can exist in multiple minds at the same time, and that individual human minds are fragmentary aspects (though not parts) of a single infinite thinking substance. These, too, are strong and counterintuitive claims.

The attraction of Spinoza's approach, however, is that it at least offers, as Locke's does not, to explain *how* it can be that one thing can, in itself, be both thinking and extended. In the 350 years since Descartes wrote, many attempts to resist his denial of extended cogitative beings have taken a broadly Lockean approach, attempting to show that extended thinkers, while metaphysically puzzling, cannot be shown to be ruled out, so that empirical findings can convince us that they may or must somehow be actual. Thus, Jerry Fodor has written:

> Nobody has the slightest idea how anything material could be conscious. Nobody even knows what it would be like to have the slightest idea about how anything material could be conscious. So much for the philosophy of consciousness.[61]

If that is indeed so, then perhaps it is time to revisit what a bolder Spinozistic approach has to offer.[62]

[61] *Times Literary Supplement*, July 3, 1992.

[62] I have benefited greatly from the helpful comments of Marleen Rozemond, Olli Koistinen, Andrew Pessin, Charles Jarrett, Talia Bettcher, Amy Schmitter, and Jon Miller.

Sensation in a Malebranchean Mind

Alison Simmons

My topic is Malebranche's theory of mind and, as my title suggests, the place of sensation within it. My aim, however, is not simply to recount a long forgotten, and rather weird, account of the human mind. My aim is to explore the roles that intentionality and consciousness play in conceptions of the mind, and Malebranche provides a particularly nice case study. That is in part because his theory of mind as a whole is so strange from our point of view. (We see all things in God?!?) In thinking through such a theory, nothing can be taken for granted, and that helps to raise questions about the mind that often go unasked. It is also a good case study because recent commentators have made a rather striking claim about Malebranche's place in the history of theorizing about the mind: Malebranche, they claim, was the first (and perhaps only) philosopher in the early modern period to break decisively with the view that intentionality is a mark of the mental.[1] A striking claim, if true. I think it is untrue. To see why requires turning over a number of rocks that, I hope, reveal what *is* strikingly novel, and thought-provoking, about Malebranche's unusual account of the mind.

Why highlight sensation? In defending the claim that Malebranche rejects intentionality as a mark of the mental, commentators point to his treatment of sensation. Malebranche, they claim, draws a sharp distinction between sensation and perception: sensation is a decidedly non-intentional mental state (a *mere* sensation, if you will) while perception is an intentional mental state.[2] I think this is wrong as a reading of Malebranche's account of sensation, and, more generally, wrong as a reading of the place of intentionality in his theory of mind. Malebranche is not interested in denying that intentionality is a mark of the mental. He is out to change our understanding of the nature of intentionality. I'm not the first to point out that Malebranche has a different conception of intentionality from many of his contemporaries,[3] but I think that commentators have not fully appreciated its consequences for the nature of Malebranchean sensation, or, consequently, for the scope of intentionality in the Malebranchean mind.

The typical strategy for investigating this topic is to look first at what Malebranche has to say about sensation, note that it seems to be a non-intentional mental state, and then draw the conclusion that Malebranche rejects intentionality

[1] The clearest recent examples are Jolley (1990, 1995, 2000), and Pyle (2003), though others effectively commit him to this position, including Alquié (1974), Guéroult (1987), Lennon (1992), Nadler (1992b), Rodis-Lewis (1963), and Schmaltz (1996).

[2] If this reading is right, then Malebranche would be anticipating Thomas Reid's more famous, and very clear, distinction between sensation and perception.

[3] Nadler (1992b) makes this point quite explicitly, clearly, and convincingly. So does Pyle (2003).

J. Miller (ed.), *Topics in Early Modern Philosophy of Mind*, Studies in the History of Philosophy of Mind 9, 105–129.

as a mark of the mental. I want to reverse things, looking first at what Malebranche says about the nature of the mind generally, and then thinking through its implications for sensation. The texts look a little different when we turn things around in this way. If I'm right, Malebranche does demonstrate a commitment to mental life being essentially intentional. And if that's right, then sensation too had better be intentional. But, we will have to ask, how? And if Malebranche does draw a sharp distinction between sensation and other sorts of mental state, as I agree he does, then what does that distinction amount to if not a distinction between the non-intentional and the intentional?

I will suggest that what is special about sensation is not its relationship to intentionality, but its relationship to consciousness. Sensations stand out to consciousness in a way that other mental states do not. But consciousness, in Malebranche's estimation, is a very poor guide to the mind, and so the fact that sensations stand out wrecks epistemological havoc both for our knowledge of both mind and world. Unlike many of his contemporaries, Malebranche argues that consciousness, while providing a *special* form of access to the mind, is not all that illuminating. In this, and in his examination of the epistemology of consciousness more generally, he stands out against others in the period. In the end, Malebranche may very well be a maverick in the history of thinking about the human mind. But that status has less to do with what he takes to be the mark of the mental than with his view that consciousness provides but a very dim light on the mind.

1 Preliminaries

I have been throwing around a philosophical term that means different things to different people and, worse, is not a word that Malebranche himself uses: intentionality. Commentators freely use the term when writing about Malebranche and I will join them, but with some caution. I mean to employ the term in a maximally inclusive way. To say that a mental state is intentional is simply to say that it is *of* or *about* or *directed to* an object of some sort (be it a physical object, a mental object, a concrete object, an abstract object, a proposition, a state of affairs, or any other sort of object you fancy). It is *not* to say, more particularly, that the mental state possesses content (propositional content, informational content, conceptual content, representational content, or any other kind of content). It is not to say that the mental state is truth-apt or has satisfaction conditions. It is not to say that the mental state is capable of being directed to a non-existent object or that its object has a special sort of immanent *inesse*. I mean my use of intentionality to be neutral with respect to competing theories of intentionality, theories about its nature. As I use the term, both direct realists (who think of mental states as involving a primitive relationship between mind and world) and so-called intentionalists (who think of mental states as having some sort of truth-evaluable content that represents the world as being thus and so, whether or not the world is thus and so and,

perhaps, whether or not the world even exists) conceive thought as intentional. I believe that the claim that Malebranche rejects intentionality as a mark of the mental does not turn on reading "intentionality" in any particular technical way, though I will have to substantiate that as I go along.

It is really the nature of *non*-intentional mental states that is important for the target thesis. A *non*-intentional mental state is going to be one that is not of, about, or directed to an object in any way at all; it doesn't even purport to be of, about, or directed to an object. Non-intentional mental states, if they exist, are a bit like mental bruises: they are typically caused by objects, but they are not in turn of or about their causes. They begin and end in the mind. Candidates for non-intentional mental states include pains, tickles, orgasms, and moods. Whether these mental states are in fact non-intentional depends, of course, on your theory of intentionality and your intuitions about these states. But for fans of non-intentional mental states, these are the ones that rise up as the best candidates.

One further note on terminology: I will use "thought" and "perception" interchangeably. This is a jarring interchange for 21st century readers, but for the early moderns both terms are used generically to pick out a variety of mental states. What we might describe as a *thought* about some mathematical proposition, they are just as likely to describe as an intellectual perception. What we might describe of as a *perception* of a pink flower, they are just as likely to describe as a sensory thought. It is not that they fail to notice that there are great differences between thinking about abstract matters and seeing concrete particulars; it's just that the difference is not systematically reflected in the use of the terms "thought" and "perception." Unless I indicate otherwise, then, these terms are simply generic terms for mental states.

2 Malebranche the Maverick

Let's start by getting the target thesis in place. Nick Jolley offers a clear statement of it:

> In general, seventeenth-century philosophers seem to have assumed that intentionality is an essential characteristic of our mental life. Malebranche is perhaps the only philosopher in the period who stands out clearly against the prevailing orthodoxy; he is committed to the thesis that there is a large class of mental items—sensations—which have no representational content.[4]

In his recent monograph, Andrew Pyle similarly points to Malebranche's treatment of sensation as evidence for his "absolute denial of the thesis that intentionality is the mark of the mental."[5] Insofar as they agree that Malebranchean sensations are non-intentional mental states, Tom Lennon, Steven Nadler, and Tad Schmaltz

[4] Jolley (1995, 128–129). See also Jolley (1990, 60; 2000, 31).

[5] Pyle (2003, 61).

commit Malebranche to a similar position.[6] French commentators say much the same.[7] The emerging consensus, then, is that Malebranche rejects intentionality as a mark of the mental, and that sensations provide the evidence.

So what's the evidence that Malebranchean sensations are non-intentional? There appear to be three sources. First, Malebranche insists over and over that sensations are just modifications of the mind. Here's an example:

> All the sensations of which we are capable could subsist without there being any object outside us. Their being contains no necessary relation to the bodies that seem to cause them, as will be proved elsewhere, and they are nothing other than the soul modified in this or that fashion; consequently, they are properly *modifications* of the soul.[8]

The force of the restriction "nothing other than" and elsewhere "merely" in classifying sensations as modifications of the mind seems to be that sensations do not represent anything in extramental reality, and (so?) are not of or about or directed to, objects outside the mind. They are purely subjective mental states, as Schmaltz puts it, exhausted by their intrinsic phenomenological features.[9]

If the ring of the text isn't enough to convince, consider Malebranche's account of secondary qualities. If secondary quality sensations are going to be intentionally related to anything outside the mind, then presumably they are going to be intentionally related to secondary qualities: sensations of color will put us into intentional contact with colors, sensations of odor will put us into intentional contact with odors, and so on. But, the argument goes, Malebranche is a sensationalist about secondary qualities; that is, secondary qualities like color and odor are themselves nothing but sensations in the mind. This is true. Malebranches writes explicitly to Arnauld that he "learned from Descartes that color, heat and pain are only modalities of the soul."[10] Here's a more elaborate statement:

> Our eyes represent colors to us on the surface of bodies and light in the air and in the sun; our ears make us hear sounds as if spread out through the air and in the resounding bodies; and if we believe what the other senses report, heat will be in fire, sweetness will be in sugar, musk will have an odor, and all the sensible qualities will be in the bodies that seem to exude or diffuse them. Yet it is certain … that all these qualities do not exist outside the soul that perceive them.[11]

If secondary qualities are themselves just sensations in the mind, then there does not seem to be anything outside the mind for those sensations to be intentionally

---

[6] See Lennon (1992, 64), Nadler (1992b, 74), and Schmaltz (1996, 94, 117).

[7] See, for example, Alquié (1974, 152–159), Guéroult (1987, 20), and Rodis-Lewis (1963, 103).

[8] *Search* I.i.1, OCM I 42–43/LO 3. See also *Search* III–II.v, OCM I 433/LO 228.

[9] Schmaltz (1996, 94, 117–118). Schmaltz offers this explicitly as an account of Malebranche's claim that sensations are "nothing but modifications of the mind." Alquié (1974) says much the same, 505–506.

[10] *Trois Lettres*, OCM VI, 201. Malebranche may be misreading Descartes, but he makes his own position on secondary qualities perfectly clear here. For an excellent discussion of Malebranche's position on secondary qualities, see Schmaltz (1995).

[11] *Elucidations* 6, OCM III 55–56/LO 569.

relating us to. And so they must be non-intentional mental states.[12] Locutions like "sensation of red" and "sensation of sweetness" must be employing descriptive genitives that simply tell us which qualitative kind of sensation we are having, rather than objective genitives telling us what feature of the world the sensation is directing us to.

But perhaps the most decisive evidence that Malebranchean sensations are non-intentional comes from what I will call Malebranche's "duplex theory" of sensory perception. Malebranche insists that sensory perception always involves two things: a *sensation* and a *pure idea*. The ideas in question here are intelligible ideas in God's mind (akin to Platonic ideas, they are eternal, immutable, infinite, universal, necessary, etc.).[13] Here's a sample passage:

> When we perceive something sensible, one finds in our perception *sensation sentiment* and pure *idea*. The sensation is a modification of our soul, and it is God who causes it in us … As for the idea that is found together with the sensation, it is in God and we see it because it pleases God to reveal it to us.[14]

Commentators routinely read "pure idea" as elliptical for "pure perception of an idea," ideas being the objects of our pure perceptions. (Pure perceptions are intellectual perceptions.) Malebranchean sensory perception, then, appears to be a mélange of two distinct and heterogenous mental states: a sensation and a pure (or intellectual) perception of an idea in God's mind. The sensation is wholly non-intentional but the pure perception is intentional because it is directed to an object outside the human mind, viz., an idea in God's mind. As Schmaltz puts is: "Sensations … bear a causal relation to divine ideas, but they lack the sort of direct cognitive relation that is characteristic of pure perceptions."[15] And Jolley writes: "although sensations may occur in conjunction with perceptions of ideas which are intentional, in themselves they are not intentional."[16] The two together constitute a sensory perception. On this view, the presence of a sensation is what makes the overall experience properly *sensory* and it phenomenologically adorns (but epistemologically shrouds) what is otherwise an intellectual perception of an idea in God.[17]

Note that on the duplex theory, sensory perception as a whole is indeed intentional, but its intentionality is secured exclusively by the perceptual/ intellectual

---

[12] Jolley (1995) employs just such an argument for the target thesis.

[13] For a taste of Malebranche's description of these ideas, see, for example, *Search* IV.xi.3, OCM II 103/LO 322 and *DM* I.vii, OCM XII 40/JS, 12.

[14] *Search* III–II.vi, OCM I 445/LO 234.

[15] Schmaltz (1996, 107–108); see also 99.

[16] Jolley (1995, 131); see also Jolley (1997, xviii).

[17] A number of French commentators offer a similar analysis of sensory perception without the explicit claim that the sensory component is non-intentional, but with the explicit claim that the sensation is joined to an intellectual perception of an idea, so that there are two distinct mental states at work only one of which is a perception of an idea. See Alquié (1974, 505), Elungu (1973, 127), and Rodis-Lewis (1963, 103, 139).

component to which sensations have attached themselves like so many mental barnacles. Of course sense perceptual experience doesn't *feel* like a complex of phenomenally impressive but non-intentional sensations and intellectual perceptions of abstracta. But like many early moderns, Malebranche maintains that there is something confused about sensory perception. As defenders of the target thesis understand it, the sensation and the perception get mixed up, or literally confused. So what happens when I have the experience of sensorily perceiving a red circle is that "a sensation of red occurs in conjunction with the perception of an idea (a geometrical concept) in such a way that I take my experience to be of a red, circular body."[18] (How to cash out the "in such a way" is, of course, the million dollar question. In fairness, though, this is a difficult thing to account for on any interpretation of Malebranche.) If sensations are simply adding a bit of phenomenological panache to otherwise intellectual perceptions, then it does seem that in and of themselves, they are non-intentional, and so we would indeed have to conclude that Malebranche gives up intentionality as a mark of the mental.

The case for my opponents looks pretty solid. I nevertheless think it is wrong. To see why it is wrong, we have to back up a bit and look at some features of Malebranche's treatment of the mind more generally, in particular what he has to say about intentionality and consciousness, and then return to ask how sensations and sensory perception fit into that account.

3 Malebranche on Intentionality

The place to turn for Malebranche's view about intentionality is his (in)famous debate with Arnauld over the nature and status of ideas. In recent Anglo-American commentaries, the debate has been cast as a debate between direct and indirect realism: Arnauld is supposed to be the progressive direct realist; Malebranche is the classic skepticism-inviting indirect realist (according to whom ideas are immediate objects of perception that mediate our access to particular physical objects by serving as representational proxies for them). While there is something to this portrayal, I think it is misleading and loads the dice in Arnauld's favor. After all, *each* of them appeals in one way or another to representative ideas in his account of cognition; and *each* accuses the other of some sort of skepticism-inviting indirectness in virtue of his peculiar use of these representative ideas. There is

[18] Jolley (2000, 40). Nadler offers a similar account: "What happens in sensory perception is that both of these elements—a conceiving and a sensing, each of which can otherwise occur by themselves—are present … One can perform an act of conceiving with one's eyes closed, and thereby intellectually apprehend a pure idea of extension undistracted by any (visual) sensations. When one opens one's eyes, the act of conceiving, which formerly took place by itself, now becomes, along with the onrushing flood of sensations, an element in our perceptual consciousness of the objects in the world around us" (Nadler 1994a, 199).

more going on in the debate than a dispute about the (im)mediacy of cognition. What is most fundamentally at issue is the nature of intentionality.[19]

Both Arnauld and Malebranche repeatedly say that *thought is always thought of something* and that *to think of nothing is not to think*. Here's Arnauld:

> Since it is clear *that I think* it is also clear that I think of something, because thought is essentially thus.[20]

And Malebranche:

> To see nothing is not to see; to think of nothing is not to think … Properly speaking, this is the first principle of all our knowledge.[21]

Both Arnauld and Malebranche at least *seem* to be committed to the view that thought is essentially intentional: thought essentially has an object of some sort. What is at issue between them is what it is that makes a thought be a thought *of something*. And that is what interests me in the debate. I do not think that what I will have to say here is in any way controversial, but I am going to draw some consequences from it that will be controversial, and so it is worth getting clear on the basics.

Arnauld holds what I will call an *intrinsic* conception of intentionality.[22] The idea here is that mental states are Janus-faced: they have formal being and objective being. They have formal being insofar as they are actual modifications of an actually existing human mind, that is, insofar as they are acts of perceiving (or, equivalently, thinking); and they have objective being insofar as they are themselves representations of actual or possible things. My visual perception of a poodle is at once an act of visual perceiving and a representation of a poodle; it is, we might say, a visual-perception-of-a-poodle. Mental states are intentional *insofar as they have objective being*, i.e., insofar as they are themselves representations of actual or possible objects. Moreover, mental states have objective being by their very nature, or essentially. Arnauld writes: "every perception is essentially representative of something."[23] In his view of intentionality Arnauld is drawing directly on Descartes.[24]

---

[19] Again, I am by no means the first to note this. Nadler (1992b) and Pyle (2003) offer extensive discussions of this aspect of the debate.

[20] *VFI* 6.

[21] *Search* IV.xi.3, OCM II 99/LO 320. See also *Search* VI-II.vi, OCM II 372/LO 481; *DM* I.4, OCM XII 35/JS 8; *Réponse* X, OCM VI 84; and *Trois Lettres*, OCM VI 202.

[22] Nadler (1992b) calls it a *content* conception of intentionality.

[23] *VFI* 6, 52; see also *VFI* 2, 22. What is more, as Pyle rightly points out, Arnauld thinks this is about as far as one can go in giving an account of intentionality: "we would only confuse and bedazzle ourselves if we tried to discover how the perception of an object can be in us … Since it is the nature of the mind to perceive objects … it is ridiculous to ask whence it arises that our mind perceive objects" (VFI 7). See Pyle (2003, 79).

[24] AT VII 42.

On this view, no actually existing object (or object substitute) need be present independent of the perception as a necessary condition for that perception, that is in order for the mind to perceive or think of it. The Janus-faced modification of the mind by itself is enough. The object-directedness of the perception is accounted for by intrinsic features of the perception or thought.

One important consequence of all this is that in the Arnauldian context *perception* and *idea* are one and the same thing ontologically, viz., a psychological state or modification of the mind. As he puts it:

> I take the perception and the idea to be the same thing. It must nevertheless be noted that this thing, although single, has two relations: one to the soul that it modifies and another to the thing perceived inasmuch as it is objectively in the soul. The word 'perception' marks more directly the first relation and 'idea' the second.[25]

Considered as a perception, the modification of mind has formal being; considered as an idea, it has objective being. This Janus-faced characterization of thought is precisely what Malebranche challenges.

Malebranche will have nothing to do with Arnauld's essentially representational perceptions.[26] He writes about them with almost palpable disgust as if he's writing about square circles. He favors instead what I will call an *extrinsic* conception of intentionality.[27] On this view, mental states have *only formal being.* They are modifications of the mind, i.e., mental acts or operations. They are not in and of themselves representations of anything; they have no intrinsic objective being whereby the objects of thought come to exist somehow in the mental act or, indeed, in the mind of the perceiver in any way. Mental states are intentional in virtue of standing in relation to actually existing things that lie outside the mind. In this sense, intentionality is extrinsic; it is a relational property of thought. Intentionality is nevertheless no mere accidental property. Recall that for Malebranche, as for Arnauld, *all* thought is thought of something, *all* perception is perception of something. Malebranche is quite clear that thoughts or perceptions *cannot exist without actually existing objects*: "I claim that there is no perception at all when there is nothing that one can perceive."[28] Intentionality, then, is an *essential* property of thought. It's just that this essential property is itself a relational property. So while the mind may not be *intrinsically* intentional, it nevertheless is *essentially* intentional.

Thus far Malebranche is in good company with naïve realists and sense-datum theorists: in order for me to perceive something there has to be or exist something there that I perceive. Unlike most naïve realists and sense-datum theorists, who have this view about sensory perception in particular (differing in what they take the requisite existing object to be—a physical object or a sense-

[25] *VFI* 5, 20.

[26] Amy Schmitter thinks that Hume also distinguishes representation and intention. See p. 228. [Ed. note.]

[27] Nadler (1992b) calls it an *object* conception.

[28] *Réponse* X, OCM VI 84.

datum), Malebranche thinks this analysis is true of all forms of thought: what makes any thought be of or about or directed to something is for it to stand in a relation to an actually existing thing of some sort. Whether I am looking at a milkshake or thinking about the nature of circularity or, indeed, hallucinating little green men dancing on my desk, my mind must be standing in relation to some sort of existing thing:

> I am thinking of a number of things: of a number, of a circle, of a house, of such and such things, of being. Thus all those things exist, at least during the time that I think of them. Surely, when I think of a circle, of a number, of being or of infinity, of a particular finite being, I perceive realities. Because if the circle that I perceive were nothing, in thinking of it I would be thinking of nothing at all. Thus at the same time I would be thinking and not thinking at all.[29]

Ultimately these existing things or "realities" turn out to be ideas in God, but I want to put aside that complication for the moment. At present what's important is that intentionality is a *relational* rather than an intrinsic property of Malebranchean thought that depends on the existence of the two relata, weird though one of them (ideas in God) may be.[30]

It is worth pausing for a moment to consider some of the motives underlying Malebranche's commitment to this extrinsic conception of intentionality, since it commits him to an inferential leap that looks fallacious to contemporary readers (and, it should be said, it did to Arnauld as well[31]): from the claim that every perception/thought is a perception/thought *of something* to the claim that every perception is a perception *of some actually existing thing*.

First, there are theological reasons for the commitment, the most important of which is that Malebranche is committed to the Augustinian doctrine that the human mind is not a light unto itself, but must be illuminated from without by God: "only He can enlighten us, by representing everything to us."[32] It is a testament to our utter dependence on God in all things that even for thought we depend on him to provide an object. As Malebranche sees it, in thought we are (cognitively) united to God.

Second, and more interesting for philosophical purposes, there are anti-skeptical and anti-psychologistic motivations at play. Malebranche worries that if our perceptions are themselves intrinsically representational then human cognition is going to traffic in private psychological representations in the mind of each cognizer. If human cognition traffics in private psychological representations in

---

[29] *DM* I.4, OCM XII 35/JS 8. See also *Search* III–II.i.1, LO I 414/LO 217.

[30] Malebranche does not have a lot more to say about this relation itself, though he does have a good deal to say about the conditions that make it possible. To be perceived, an object must be capable of causally affecting the mind, and to causally affect the mind it must be intelligible. Ideas in God fit the bill. Beyond that, explanation seems to give out: "the nature of the soul is to perceive what touches it" (*DM* V.v, OCM XII 115/JS 76).

[31] See *VFI* 4.

[32] *Search* III–II.vi, OCM I 447/LO 235. See also *Search* IV.xi.3, OCM II 97–98/LO 319.

the mind of each cognizer, then, he supposes unnecessarily but unoriginally, we must really all be cognizing numerically distinct things, viz., our own representations. And if we are cognizing numerically distinct things, we have no guarantee of either *intersubjective agreement* (that we are all thinking the same thing when we think, for example, about the Pythagorean Theorem) or *objectivity* (that what we are thinking about corresponds to something outside our own minds). Representational perceptions thus invite dangerous forms of psychologism and skepticism.[33]

This result, Malebranche protests, is unacceptable in a theory of cognition. Whether you and I are looking at milkshake or contemplating the Pythagorean Theorum, it had better turn out that we are looking at and thinking about *the very same things*,[34] and that those things lie *outside either of our minds*.[35] These are the non-negotiable facts of cognition that any theory must account for. To Malebranche's mind the best way to account for them is to distinguish sharply the acts and objects of cognition. Our acts of seeing and thinking are numerically distinct—they are subjective psychological modifications of your and my mind, but the objects we are seeing and thinking about are not:

> The perception I have of intelligible extension belongs to me, it is a modification of my mind. It is I who perceives this extension. But this extension that I perceive is not a modification of my mind. For I am well aware that it is not myself that I see when I think of infinite spaces, of a circle, of a square, of a cube, when I look at this room, when I turn my eyes to the sky … The perception I have of extension could not exist without me. It is therefore a modification of my mind. But the extension I see subsists without me. For you can think of it without my thinking of it, you and everyone else.[36]

Malebranche seems to think that a theory of cognition that employs psychological representations sacrifices the non-negotiables. In his snarkier moments, he resorts to crude insult: Arnauld, he charges, must be unable to love a woman, for to love a woman you have to see a woman, but all Arnauld sees are *representations* of women that are, in fact, modifications of *himself*. And so Arnauld's theory of cognition renders him capable of loving only himself![37]

---

[33] Pyle (2003, 85–89), makes a similar point.

[34] In the domain of understanding eternal truths, such as that twice two is four, the Chinese, he insists, "see the same truths as I do" (*Elucidations* X, OCM III 129/LO 613). As for sensory perception: "A thousand people can see one and the same column, and I mean numerically the same" (*Réponse* XIII, OCM VI 98).

[35] In part he argues from the phenomenology of the experience: "When I see a centaur, I note in myself two things. The first is that I see it; the second is that I am aware that I see it. I see it, but as being distinct from myself. It is therefore not a modification of my substance" (*Réponse* VI.ix, OCM VI 60). In part he argues that if the objects are internal to the mind, we face skeptical worries: "If our modes of mind were representational, how would we know that things correspond to our ideas?" (*Search* IV.xi.3, OCM II 99/LO 320).

[36] *DM* I.ix, OCM XII 45/JS 16.

[37] *Réponse* IX, OCM VI 77–78. Granted, Arnauld had provoked this response by charging Malebranche with the view that women who admire their beauty actually see God when they look in the mirror, "since the face they see is not theirs but an intelligible face that resembles theirs in God" (reproduced in *Réponse*, OCM VI 76).

Arnauld will in effect argue back that that the representational ideas in question are not in fact *objects* of cognition but only *vehicles* of cognition: I don't *see* my idea of a milkshake; I *have* an idea of a milkshake. And in having it I thereby see *the milkshake*, assuming it exists (else I only seem to see it). Similarly, I'm not *thinking about* some psychological representation of the Pythagorean Theorum flitting around in my own mind; I *have* that representation in my mind, and in having it I thereby think about an eternal truth that is utterly unaffected by the vicissitudes of my mental life. It is Malebranche, Arnauld will argue, who is walking into the skeptic's trap, for, as I'll mention in a moment, the non-psychological objects he thinks we are all immediate acquainted with in sensory perception are not, in fact, milkshakes and doughnuts but representational ideas in the mind of God.[38] There is a lot of willful misunderstanding in the debate between Arnauld and Malebranche. What I want to highlight is that both are motivated to avoid skepticism, and that each seems to think that his account of what makes a mental state be of or directed to an object best skirts the danger.

So let's face the strange twists in Malebranche's otherwise unremarkable account of the mind's extrinsic intentionality. First, the actually existing objects to which most of our mental acts are directed are not physical objects but *ideas in God's mind*. It would take us too far afield to rehearse Malebranche's many arguments for this strikingly bizarre claim, but one crucial argument rests on the idea that for an extramental thing to be perceived by the mind it has to be able to act on the mind, and God alone, Malebranche argues, can act on the mind.[39] What is of interest to me is the consequence that the distinction between *perception* and *idea* in Malebranche is ontological: perceptions are modifications of the human mind; ideas are ontologically distinct objects of those perceptions that exist in the mind of God and that are available alike to all minds. This is a distinction that Malebranche presses insistently and persistently in his writing, and it is understandable that he does so, since Cartesians like Arnauld and Descartes himself are committed to the view that perceptions and ideas are one of the same thing, viz., modifications of the human mind, considered in two different ways, formally and objectively.

Second, the ideas in God's mind are representational entities of some sort. Among the things they represent are concrete particulars, so that we perceive milkshakes and doughnuts only indirectly by way of perceiving ideas in God that represent them. Now lest we think that God is walking around thinking about milkshakes and doughnuts, or that Malebranche is a classic indirect realist with just the one minor (!) twist that the representational ideas are in God rather than us, we should note that the *way* in which God's ideas represent concrete particulars is itself unusual. Ideas in God are not *copies* or *images* of concrete particulars. They are (again, more like Platonic ideas) the universal *models* or *archetypes* for concrete particulars, and each is a model or archetype for infinitely

[38] *VFI* 4.

[39] For good summaries of his arguments see Lennon (1992) and Pyle (2003, Ch. 3).

many possible concrete particulars. In the case of physical things, Malebranche's official position appears to be that there is really just *one* idea in God that represents, or serves as the archetype for, all actual and possible bodies, viz., the idea of intelligible extension. This idea is *the* idea or archetype "of the material world in which we live, and of an infinity of other possible worlds." [40] Just how this idea represents all actual and possible concrete particulars is a matter of interpretive debate,[41] but it is reasonably clear that Malebranche thinks that whether we are looking at a milkshake or doughnut, or for that matter thinking about circularity, it is really one and the same object we are perceiving, viz., the idea of intelligible extension, which idea seems to represent all of these concrete particulars and essences, and which serves as the basis for creation.[42]

These peculiarities of Malebranche's account are not what interest me except insofar as they illustrate that locutions like "thought of a circle" and "perception of a milkshake" pick out rather different things in the Cartesian and Malebranchean contexts. In the Cartesian context, they pick out modifications of the human mind with some sort of intrinsic representational content. In the Malebranchean context, they pick out relations between modifications of the human mind and ontologically distinct representational ideas in God.

So far, I don't think I've said anything terribly controversial (interpretively, anyway), but there is an important consequence that I think is often overlooked: intentionality and representationality come apart in Malebranche in a way they do not for Descartes and Arnauld. Intentionality is a relational property of the mind's perceptions; representationality is a property of ideas in God. Intentionality is essential to the mind; representationality is not only inessential to but anathema to the (human) mind. Commentators like Jolley and Pyle float breezily between talk of intentionality and representationality. Jolley writes in the passage quoted at the start of the paper:

> Malebranche is perhaps the only philosopher in the period who stands out clearly against the prevailing orthodoxy that intentionality is an essential characteristic of our mental life;

[40] *DM* II.ii, OCM XII 51/JS 21. See also *Elucidations* X, OCM III 153–154/LO 627 and *Réponse* VI, OCM VI 61. There has been a great deal of interpretive discussion concerning the question whether there are in God ideas that correspond to every particular body (the sun, my right hand), ideas simply of the various kinds of body (horse, tree), or only a single idea of extension. Malebranche's position seems to have developed, resulting in his mature view is that there is only a single idea representative of body. For a nice recent discussion, see Reid (2003).

[41] For helpful accounts of the way in which ideas in God represent concrete particulars, see Radner (1978) and Reid (2003).

[42] Radner helpfully notes that by placing representative ideas in God, and conceiving them as archetypes or models for creation, Malebranche effectively bypasses the skeptical worries that beset Descartes about whether ideas actually correspond to things in the world. They cannot help but correspond to them, in Malebranche's system, because they are the very models for them. See Radner (1978, 61–62).

> he is committed to the thesis that there is a large class of mental items—sensations—which have no representational content.[43]

And Pyle:

> Malebranche insists on a sharp contrast between sensations, which are modes of our souls and represent nothing beyond themselves, and ideas, which are 'in' God and represent objects. His rejection of the *monde intelligible* hypothesis the view that our mental states themselves make the world intelligible by representing it thus carries with it an absolute denial of the thesis that intentionality is the mark of the mental.[44]

This is a mistake in the Malebranchean context. It's a natural mistake to make these days, since today the two notions are used more or less interchangeably.[45] But it is a mistake nevertheless.

Now one might reply that Malebranchean perceptions are intentional only insofar as the items they are related to are representational ideas in God, so that intentionality piggybacks on representationality after all. I don't think that's right. Quite apart from the peculiarity of the way in which ideas in God are supposed to represent concrete particulars, not all Malebranchean perception operates indirectly by way of ideas in God. Perception of body (in the abstract and in the concrete) does because body is not itself intelligible; they must therefore be represented to us. Ideas render bodies intelligible by representing them. Things that are intelligible in themselves, by contrast, are supposed to be perceivable in themselves without the mediation of representational ideas: "We know things by themselves and without ideas when they are intelligible by themselves, i.e., when they can act on the mind and thereby reveal themselves to it."[46] Thus we are supposed to perceive God "through a direct and immediate vision" without any mediating representational ideas.[47] Malebranche tentatively allows for the immediate perception of other intelligences, though he thinks human perceivers are barred from it in this life.[48] This unmediated form of perception is still intentional; it still has an object. And yet it is free from any form of representation.

The interpretive payoff of all this is that while in a Cartesian context a mental state's being non-representational would entail its being non-intentional, in the Malebranchean context it does not. And so to establish that sensations are non-intentional it is not enough to establish that they are non-representational. And it had better not be enough, for *no* state of the human mind is representational on Malebranche's view. Not sensations. Not purely intellectual perceptions. That's

---

[43] Jolley (1995, 128–129).

[44] Pyle (2003, 61).

[45] Charles Siewert's entry on consciousness and intentionality in the *Stanford Encyclopedia of Philosophy* thus notes: "Intentionality includes, and is sometimes take to be equivalent to, what is called 'mental representation.'"

[46] *Search* III-ii.7.1, OCM I 448/LO 236.

[47] *Search* III-ii.7.2, OCM I 449/LO 237.

[48] See, for example, *Search* III–II.i.1, OCM I 415-417/LO 218 and *Search* III–II.vii.1, OCM I 448/LO 236.

precisely the point against Arnauld: *no mode of the human mind is representational*. If lack of representationality entails lack of intentionality, then Malebranche is committed to the very extreme view that the mind is exhibits *no intentionality at all*. And that is not a position commentators would want to saddle him with. To show that sensations are non-intentional, then, one has to establish that they are not directed to any object (be it an idea in God or, more generally, anything intelligible).

4 Malebranche on Consciousness

Let's turn to consciousness. Like Arnauld and Descartes, Malebranche will say that consciousness (*conscience*) amounts to an immediate awareness of one's mental states. More precisely, he describes consciousness as a kind of "inner sentiment" of the modifications of one's own mind: "by 'consciousness' *conscience* I understand inner sentiment *sentiment intétieur*."[49] And what we inner sentiment of are the modifications of our own mind: "the inner sentiment I have of myself teaches me that I am, that I think, that I will, that I feel, that I suffer, etc."[50] Moreover, consciousness, on Malebranche's view, seems to be a property of first-order mental states themselves; it is not a higher-order affair involving perceptions of perceptions, thoughts of thoughts, or any kind of deliberate introspection.

Consciousness appears to be a property of *all* mental states for Malebranche, regardless of their type:

> by the words *thought*, *manner of thinking*, or *modification of the soul*, I generally understand all those things that cannot be in the soul without the soul perceiving them through the inner sentiment it has of itself—such as its own sensations, its imaginations, its pure intellections, or simply its conceptions, as well as its passions and natural inclinations.[51]

I want to stress the all-inclusiveness here. *All* types of mental state, *including pure intellectual perceptions*, are conscious; that is, we have an inner sentiment of all of them. Still, there are key differences:

> The three ways in which the soul perceives, namely by the senses, by the imagination, and by the mind i.e., the pure intellect do not affect it equally, and as a result, it does not pay equal attention to what it perceives by means of each of them because it attends greatly to what affects it greatly, and little to what affects it little.[52]

---

[49] *Réponse* VI, OCM IV 56. See also *Search* III–I.i.1, OCM I 382/LO 198 and *Search* III–II.vii.1, 4, OCM I 448, 453/LO 236, 239.

[50] *DM* III.vii, OCM XII 67/JS 34.

[51] *Search* III–II.i.1, OCM I 415/LO 218.

[52] *Search* I.xviii.1, OCM I 177/LO 79. See also *Search* I.i, OCM 42/LO 2; *Search* I.xii, OCM I 140-141/LO 59; *Search* IV.xi.3, OCM II 102/LO 322; and *CC* III, OCM IV 75–76.

In other words, we are conscious of all our mental states, but we are more conscious of some than others. Sensations are more affective than pure perceptions, and so they stand out more to consciousness; they have, we might say, a more pronounced phenomenology. The argument for this claim amounts to a simple appeal to ordinary experience: "Nobody can doubt that the littlest sensory pain is more present to the mind and renders it more attentive than meditation on something of much greater consequence."[53]

Just as Malebranche distinguishes *ideas* from perception, so he distinguishes *consciousness* from perception. Both consciousness and perception belong properly to the mind, and both are functions of the human understanding in particular (as opposed to the will). Both, therefore, are cognitive functions, but they have different objects: while perception is directed outward toward ideas in God (and other intelligibilia), consciousness is directed inward to the mind's own modifications.[54] They also, Malebranche maintains, provide a different *kind* of access to their objects: consciousness gives us only a subjective access to our own mental states—in being conscious of a pain or a thought I *feel* it (hence "inner sentiment") and *only I* feel it; perception, by contrast, gives us a more objective access to ideas and other intelligibilia—in perceiving we observe things as objects distinct from us, objects we can perceive in different ways, from different angles, along with other people, etc.[55] As two forms of cognition, consciousness and perception have a different *epistemic* standing. Consciousness provides only an "obscure and confused" acquaintance with our mental states, while perception is capable of providing "clear and distinct" systematic knowledge (*scientia*) of its object. And here Malebranche departs from many of his contemporaries. The key to the difference is intelligibility. Ideas render the things they represent to the mind intelligible, and so when we perceive something by way of an idea we are in a position to understand its nature or essence, derive the various properties of which it is capable from that nature, and determine the various relations that hold among those properties; it is perception the idea of extension, for example, that yields geometry. Thus he writes of our perception of body:

> the knowledge that we have of bodies and their properties is quite perfect; that is, the idea that we have of extension suffices to make us know all the properties of which extension is capable, and we could desire to have no more distinct or fruitful and idea of extension, of shapes, and of motions than the one God gives us.[56]

---

[53] *Search* I.xviii.1, OCM I 177/LO 79.

[54] *Search* I.i.1, OCM I 43/LO 3. Consciousness, then, may give us another example of an intentional mental state that does not involve representation. It is, in fact, a limiting case of Malebranchean intentionality, for the mental state is both intentional act and object at once.

[55] Tad Schmaltz, self-consciously invoking Thomas Nagel, describes the difference between consciousness and perception in terms of "subjective" and "objective" access throughout his discussion in Schmaltz 1996, and I am borrowing these terms from him.

[56] *Search* III–II.vii.3, OCM I 450/LO 237. Note that it is our perception of the *nature* of body in general, and of its possible modifications, that is capable of perfection. Like all the other early

Because consciousness operates without an idea to render its object intelligible (and because its object is not intelligible in itself), it offers us only a kind of blind, or at least highly myopic, contact with the mind:

> The inner sentiment I have of myself makes me aware that I am, that I think, that I will, that I feel, that I suffer, etc., but it does not at all make me know what I am, the nature of my thought, of my will, of my feelings, of my passions, of my pain, nor the relations all these things have to each other.[57]

A genuine a priori science of the mind, on a par with geometry, is out of the question. Instead our acquaintance with the mind is piecemeal and empirical: we learn about it by one modification of the mind at a time, as they crop up in our own mind.[58]

Putting intentionality and consciousness together, there seem to be three components to any Malebranchean mental state: (a) a perception, which is a modification of the mind; (b) an idea, or some other intelligible, which serves as the object of that perception; and (c) consciousness, which is a kind of subjective awareness of the perception.[59]

5 Malebranche on Sensations

So how do sensations fit into the Malebranchean mind? First, it is worth noting that terms like "sensation" and "sensing" are ambiguous in Malebranche's texts. Sometimes they refer to the inner sentiment or consciousness we have of our own mental states. I *sense* the various modifications of my mind, that is I *am conscious* of them. Thus:

> It is the understanding that perceives modifications of the soul, or that *senses* them.[60]
> I *sense* that I think, that I want, that I desire, that I suffer.[61]

Sensing here is being distinguished from *perceiving by way of ideas*:

---

moderns, Malebranche will maintain that our perception of *particular existing* bodies and their modifications is epistemologically problematic.

[57] *DM* III.vii, OCM XII 67/JS 34. See also *Search* III–II.vii.4, OCM I 451–453/LO 237–239; *Search* IV.xi.3, OCM II 98/LO 320; *Elucidations* XI, OCM III 163–171/LO 633–638; *Réponse* VI, OCM VI 55; *Réponse* X, OCM VI 86; and *Réponse* XXIII, OCM VI 162.

[58] This epistemic difference is, of course, what provides the basis for Malebranche's argument against Descartes' claim that mind is better known then body in *Search* III–II.vii and *Elucidations* XI.

[59] For present purposes I am leaving the volitional side of the Malebranchean mind aside, but clearly sensations are not going to find their home there.

[60] *Search* I.i, OCM I 43/LO 3, italics mine.

[61] *Réponse* XXIII, OCM VI 163, italics mine. See also *Réponse* VI, OCM VI 57 and *Search* III–II.vii.4, OCM I 451/LO 237.

> In order to *know* one needs ideas that are different from the modifications of the mind. But ideas are not necessary in order to *sense* what occurs in oneself.[62]

This is just the distinction between consciousness and perception I discussed in the previous section. And so although sensing, in this sense, is a more immediate form of cognition than perceiving by way of ideas, it is an epistemically inferior form of cognition—a sort of mere subjective acquaintance that does not afford anything like the systematic knowledge that perceiving an idea affords us. Call this use of sensation language, sense_1.

Other times Malebranche uses the language of "sensation" and "sensing" to refer to one of the various kinds of modification of mind:

> Sensations are nothing but ways of being of the mind, and it is for this reason that I call them modifications of the mind.[63]

In this usage, call it sense_2, sensation is being contrasted with *pure or intellectual perception*. Putting the two together: we sense_1 our sensations_2; that is, we are conscious of our sensations.

Third and finally, Malebranche sometimes uses the expressions "sensation" (*sensation*) and "sentiment" (*sentiment*) to refer to secondary qualities like colors, sounds, odors, flavors, hot, cold, pain, etc. Thus:

> Although we hear sounds as if spread out in the air, it does not follow that they are there. They are really found only in the soul, for they are simply *sensations sentiments* which affect it, modifications which belong to it.[64]

Call this use sensation_3. Here sensation is being contrasted with *modifications of the body*. Malebranche's point is that secondary qualities (sensations_3) are modifications of the mind (sensations_2) not modifications of the body.

The ambiguity is understandable. Sensation_1 and sensation_2 are easily confused because sensations_2 are characterized as modifications that greatly affect the mind, making them especially sensible_1 to the mind, more sensible_1 than pure perceptions. We still do sense_1 pure perceptions, but, Malebranche suggests, this is a less phenomenologically impressive affair. Sensation_2 and sensation_3 are rightly collapsed since Malebranche's official position is that secondary qualities like color and odor (sensations_3) just *are* modifications of mind (sensations_2). And the relationship between sensation_1 and sensation_3 is a tight one too, for since secondary qualities (sensation_3) just are modifications of mind (sensations_2), they are especially present to consciousness (sensation_1). Putting the three together: to sense a color or feel a pain is to sense_1 a sensation_2, or to be conscious of a modification of mind.

All of this is important, for Malebranche is one of the few very *clear* cases of a sensationalist about secondary qualities: colors and smells are modifications of

---

[62] *Réponse* V, OCM VI 54. See also *Réponse* X, OCM VI 86.

[63] *Search* I.i, OCM I 42/LO 2.

[64] *DM* III.xii, OCM 77/JS 43, italics mine. See also *CC* III, OCM IV 75 and *Search* III–II.vii.4, OCM I 452/LO 238.

the mind.[65] And this has clear implications for the cognition of secondary qualities: strictly speaking, colors and smells are not *perceived* but only *sensed*$_1$. Primary qualities like size and shape, by contrast, are perceived. There is, then, for perhaps the first time, a very deep and clear distinction between *sensing secondary qualities* and *perceiving primary qualities*.[66] It is precisely this admittedly dramatic distinction, I think, that defenders of the target view are cottoning onto in their defense of sensations as the first genuinely non-intentional mental states. But is that right?

So far it looks like sensation, in all its senses, lives entirely in the human mind. And here is where Jolley and the others call it a day: sensations are modifications of mind that we sense$_1$ and that (therefore?) have no object outside themselves. It's the second part of this claim that I take issue with. Malebranche's general account of the mind seems to commit him to the position that sensations, like all modifications of the mind, are intentional. But is that possible? Yes. It is possible because sensations$_2$ are Janus-faced. Not in the Cartesian sense of having formal and objective being, but insofar as we can think of these modifications of mind as *objects* of consciousness or as *acts* of perception and, as acts of perception, directed intentionally to ideas in God's mind. The same is true of pure perceptions: they too are modifications of mind that are at once *objects* of consciousness and *acts* of perception directed intentionally to ideas in God. The difference is that pure perceptions are relatively unremarkable objects of consciousness. Sensations, by contrast, are rather dramatic objects of consciousness. Perhaps for that reason, commentators tend to overlook their role as *acts of perception*.[67]

So what makes me think that sensations$_2$ are, like pure perceptions, acts of perception? For one thing, it is the only way to fit them into Malebranche's mental taxonomy. As Malebranche depicts it, the human mind has two chief faculties: the

---

[65] While Descartes twice suggests that secondary qualities like color are mere sensations (*Sixth Replies*, AT VII 440 and *Principles* I.68, AT VIII-A 33), he more often suggests that they are some sort of disposition, where that means either an arrangement of primary qualities in a body (*Principles* IV.198-199, AT VIII-A 323; *Fourth Replies*, AT VII 254; and letter to Regius, May 1641, AT III 372–373), a physical disposition of such an arrangement to affect the incident light in distinctive way (*Principles* III.32 AT VIII-A 184–185; *Principles* IV.131, AT VIII-A 274; *Dioptrics* AT VI 84–85 and 91–92; and *Meteors* AT VI 335), or a disposition of such an arrangement to produce certain sensations in the human mind (*Principles* IV.29, AT VIII-A 229; *To Mersenne*, 15 November 1638, AT II 440–441; and *To Reneri for Pollot*, April or May 1638, AT II 44). (The more famous passages from the *Principles* IV.191 and IV.198 that make Descartes sound like a Lockean dispositionalist, according to which secondary qualities are dispositions of the primary qualities of bodies to produce certain sensations in us, actually turn up in the amended French edition of the text by Picot.)

[66] This is, in fact, an even deeper distinction than anything we find in Reid, who resists identifying secondary qualities with sensations, and who therefore will not say that we only sense secondary qualities while we perceive primary qualities.

[67] Although I think that too few commentators recognize this, there are notable exceptions. Daisie Radner and Jasper Reid explicitly recognize sensation as a form of perception directed to ideas (Radner 1978, 1994; Reid 2003).

understanding and the will. The understanding is the faculty "of receiving various ideas, that is, of perceiving various things,"[68] while the will is the faculty of receiving inclinations or willing various things. All mental states are modifications of one of these two chief faculties. There is simply no room in this framework for non-intentional sensations that are neither ways of perceiving or ways of willing. Sensations are clearly not modifications of the will. They must therefore be modifications of the understanding. And that is just what Malebranche says. The understanding is capable of three different *ways of perceiving ideas*: sensory perception, imagination, and purely intellectual perception.[69] My suggestion is that sensation is just the same thing as sensory perception; it is one of the three ways of perceiving ideas. Does that fit with the texts? Yes. Malebranche routinely classifies sensation as a *kind of perception* alongside pure perception:

> The soul perceives by the senses only sensible and gross objects … these sorts of perceptions are called *feelings sentiments* or *sensations sensations*.[70]

Even when he asserts that secondary qualities are sensations, he sometimes identifies them alternatively as perceptions:

> Colors are not spread out on objects; they are only sensations, perceptions, or modifications of the soul.[71]

So given the structure of the Malebranchean mind and the texts, it looks to me like the best way to understand sensation$_2$, if we can, is as a kind of perception. That is, sensations are themselves perceptual acts of the mind. "Sensory perception" signifies not a complex of (non-intentional) sensations plus (intentional) pure perceptions, but a distinctive kind of (intentional) perception.

6 But What About Those Arguments?

So what about all those arguments in support of the view that sensations are non-intentional mental states that constitute a distinct component of sensory perception? The first argument went as follows: look, Malebranche says that sensations are nothing but modifications of mind. Why would he say that if not to *contrast* them with perceptions (also modifications of the mind) that reach outside the mind? The claim that sensations are only modifications is indeed doing a lot of work for Malebranche, but it is not, I think, the work of distinguishing sensations as non-intentional mental states from their intentional counterparts, perceptions. Sometimes

[68] *Search* I.i.1, OCM I 41/ LO 2.

[69] *Search* I.i.1, OCM I 41-44/ LO 2–3. See also *Search* I.iv.1, OCM I 66-67/LO 16–17. Malebranche sometimes omits imagination, since it is a faculty that derives from sensory perception.

[70] *Search* I.iv.1, OCM I 67/LO 17.

[71] *CC* III, OCM IV 75–76. See also *Réponse* XIII, OCM VI 98.

he is making a point about secondary qualities (sensations$_3$): colors, for example, are nothing but modifications of the mind (sensations$_2$), *as opposed to modifications of the body*.[72] Other times he is indeed making a point not about secondary qualities but about sensations themselves (sensations$_2$), but the point is that they are only modifications of mind *as opposed to representational entities*. As such, they are being distinguished not from pure perceptions (which aren't representational entities either) but from representational ideas–both Malebranchean representational ideas which reside in God and Arnauldian representational ideas which reside in the human mind. Consider for example the following passage:

> All the sensations and all the passions of the soul represent nothing outside of the mind that resemble them and they are only modifications of which the mind is capable. But the difficulty is to know if the ideas that represent something outside the soul and which resemble them in some way, like the ideas of the sun, a house, a horse, a river, etc, are only modifications of the soul.[73]

Nothing in Malebranche's denial that sensations represent anything outside the mind tells against their being intentional, that is being directed to (representational) ideas in God.

Enter the second argument: but there simply is nothing to serve as the intentional object of sensations$_2$! Recall Jolley: "the only way in which sensations could have intentionality is if they were directed to independently existing objects—ideas in God; but on Malebranche's view they are not."[74] Let's see about that. It's true, as Jolley points out, that they are not directed to any kind of secondary qualities (or secondary quality ideas in God). After all, sensations$_2$ *constitute* secondary qualities. There are no secondary qualities out in the corporeal world and if God has secondary qualities ideas, they are not accessible to the human mind, since we know about secondary qualities only by sensing$_1$ our own sensations$_2$.

One might float the idea that sensations$_2$ are intentionally directed to themselves so that they are their own object. And in a sense, sensations$_2$ *are* their own objects, for they are the objects of consciousness (sensation$_1$). But pure perceptions are their own objects in that sense too. We're looking for an object of *perception*, not an object of consciousness.

Is the case closed then? No. There is one other option: the idea of extension in God. Now sensations$_2$ do not *represent* extension or any modification of it. Jolley and the others quite rightly insist on that. But neither do pure perceptions of extension. As I keep pointing out, *no* Malebranchean perception represents *anything*. That's the point of rejecting Arnauld's representative perceptions. But both sensations and pure perceptions are *directed to* the idea of extension; they are different *ways of perceiving* extension. They give us different lenses, if you will, on extension. One yields something like conceptual thought about the nature of extension and its various possible modifications in general; the other yields

[72] See, for example, *Search* III–II.vii.4, OCM I 452/LO 238.
[73] *Search* III–II.v, OCM I 433/LO 228.
[74] Jolley (1995, 131).

sensory experience of particular bodies that appear to be covered in colors, redolent of smells, and hot or cold to the touch. Both, however, are fundamentally ways of perceiving the idea of extension in God. Malebranche is actually reasonably clear about that:

> When the idea of extension affects or modifies the mind with a pure perception, then the mind simply conceives this extension. But when the idea of extension touches the mind in a more lively manner, and affects it with a sensible perception, then the mind sees or senses extension. The mind sees it when this perception is a sensation of color. And the mind feels it or perceives it in a more lively manner when the perception with which the intelligible extension modifies it is a pain. For color, pain, and all the other sensations are only sensible perceptions produced in intelligences by intelligible ideas.[75]

And again:

> There are neither two kinds of extension nor two kinds of ideas that represent them. And if that extension you think of were to touch you or modify your soul by some sensation, however intelligible it is, it would appear sensible to you. It would seem hard, cold, colored, and perhaps painful; for perhaps you would attribute to it all the sensations you have … The same idea of extension can be known, imagined or sensed according to the way the divine substance containing it applies it diversely to our minds.[76]

Fundamentally, then, the senses and the intellect have the same object when they are occupied with bodies. The experience is quite different because sensations render the idea of extension sensible: color sensations make it visible, sound sensations make it audible, tactile sensations make it capable of being felt. All of these are so many ways or perceiving extension: bluely, redly, stinkily, sweetly, painfully, etc. Sensations also give the idea of extension, itself a single universal, the appearance of many particular existing bodies of determinate shapes and qualities by, as it were, illuminating different portions of it. Although the process by which the appearance of extension is supposed to transform from universal to particular is a bit mysterious, it is sufficiently clear, I think, that what is supposed to be changing is the perceptual act itself, which has the effect of changing the appearance of its object.

So what then, finally, of the argument based on Malebranche's claim that sensory perception involves two things, sensation and pure idea? If sensation just is a form of perception then what are we to make of the claim that sensory perception has two components? Here, recall, is the classic passage:

> When we perceive something sensible, one finds in our perception *sensation* and pure *idea*. The sensation is a modification of our soul, and it is God who causes it in us … As for the idea that is found together with the sensation, it is in God and we see it because it pleases God to reveal it to us.[77]

---

[75] *CC* III, OCM IV 75.

[76] *DM* II.vii, OCM XII 60–61/JS 29–30. See also *DM* V.v, OCM XII 115–116/JS 76–77.

[77] *Search* III–II.vi, OCM I 445/LO 234.

My response by now is probably obvious. Sensory perception involves two things in just the way that any perception involves two things: there is a modification of the mind (the perceptual act, viz., the sensation) and an idea in God (the object to which the perceptual act is directed, viz., the idea of extension). The sensation just *is* the perceptual act. But it is an unusually sensible$_1$ perceptual act. This is what is distinctive about sensation: it is unusually sensible$_1$ because it has a qualitative character of its own that stands out to consciousness. (Malebranche will bite the bullet and say that in having sensations the soul itself is colored, smelly, hot or cold, so that the features that make sensory perceptions phenomenologically distinctive, compared to intellectual perception, are *formal* properties of the mind.[78]) Because it stands out forcefully to consciousness, Malebranche proposes, in sensory perception we naturally mix together, or confuse, the perceptual *act* and *object*; thus, for example, we confuse sensations of brown, cold, and chocolatey flavor with the idea of extension to yield (and here is where the mystery is) a sensory perception of a tasty chocolate milkshake.[79] This may be an unlikely account of sensory perception. It is not my intention to resurrect it. What is interesting in it is the extent to which Malebranche challenges the prevailing assumptions about the roles that intentionality, representation, and consciousness all play in our mental life: intentionality and representationality come apart, and consciousness proves a poor guide to self-knowledge and a distorting influence on our perception of body.

Conclusion

As I suggested at the outset, if Malebranche is a maverick in early modern theories of mind, it is not for his rejecting intentionality as a mark of the mental. Still, his particular treatment of intentionality is unusual among his peers, and it raises some interesting questions. Malebranche conceives the mind's intentionality as an extrinsic relation between mind and extramental objects, and he divorces intentionality from representationality, removing the latter entirely from the human mind. His opponents, by contrast, take intentionality to be an intrinsic property of the mind that is accounted for by its brute ability to represent things. It is worth persistently asking ourselves which sort of view of intentionality we assume in our own theories of mind and whether intentionality is beholden to representation, as

[78] *Eludications* XI, OCM III 166/LO 634–635.

[79] There is another sense in which sensory perception could be said to be a duplex of two mental acts: it is a duplex of consciousness of a sensation and perception of the idea of extension. But again, this can be said of intellectual perception as well (it is a duplex of consciousness of the pure perception and perception of an idea), with the caveat that consciousness of a pure perception is attentively recessive, so that we do not notice it or confuse it with the perception's ideational object. This difference is a matter of degree, not kind, and it still involves no non-intentional mental states.

it is typically taken to be. Might there be mental states that are intentional but, for all that, are not representational? (And if we adopt an extrinsic model of intentionality of some sort, how do we best account not only for the difference between veridical and hallucinatory experiences, but also for the difference between sensory perception and more intellectual forms of cognition if not from difference in mental representation? Same relation to different objects? Differences in the perceiving subject?) It is not at all clear that we have yet untangled all the threads that connect these phenomena. And versions of these issues still underlie a number of current debates in the philosophy of mind (e.g., the debate between perceptual disjunctivists and intentionalists). One thing that is curious about Malebranche is that he just does not fit any of our basic categories in the theory of perception: he's not a sense-datum theorist, not an adverbialist, not an intentionalist, not a naïve realist or disjunctivist. He shares some commitments with each of them and rejects others. But it is for precisely that reason his treatment of intentionality proves thought-provoking and instructive.

If Malebranche does not depart from his contemporaries in his view about the essentiality of intentionality to the mind, he does depart from them in his view about consciousness. Like other early moderns, he brings consciousness into focus as an essential feature of the mind. And it is easy to suppose that he is like the others in thinking that consciousness provides an unmediated and so *better* access to the mind than the access perception, in its various forms, provides to extramental reality. It is true that Malebranche thinks that consciousness of our own mental states does not require ideas, whereas perception of most extramental things does. But this doesn't make it better.

In a much more systematic way than his contemporaries, Malebranche undertakes an explicit examination of the difference between knowing through consciousness and knowing through other forms of cognition (including sensory and intellectual perception); he explores, we might say, the difference between first-person and third-person access.[80] In his estimation, consciousness is an *inferior* form of cognition that provides only superficial acquaintance with our own mind—both in its particularity and in its nature. True, it gives us quite certain knowledge that we are having this or that mental state, when we do. But for all that it does not make those states, or the nature of the mind more generally, intelligible. "No matter how much effort I put into attending through consciousness to a mental state, it isn't intelligible to me."[81] Or again:

> It is true that we know well enough by our consciousness or by the inner sentiment we have of ourselves, that our soul is something important. But it could be that what we know of it may be almost nothing of what it is in itself. If one knew only of matter twenty or thirty shapes by which it can be modified, surely one would know almost nothing in comparison to what we can know by the idea that represents it. It is therefore not enough to know the soul perfectly to know only what we know by inner sentiment; since the

[80] Tad Schmaltz thematizes this point in Schmaltz (1996).

[81] *Réponse* OCM VI 162.

> consciousness that we have our ourselves may perhaps show us only the least part of ourselves.[82]

The investigation of the mind through consciousness yields fruits wholly unsuitable to anything like a science of the mind. Many of the mental states uncovered are ineffable (secondary quality sensations cannot be defined or described, but only experienced); what it takes for the mind to have these mental states remains unknown; what the precise relations are between mental states is not revealed; and, again in the case of sensations, we aren't even really sure whether what we are experiencing is properly a state of the soul or a state of body. All this does raise a set of good question: What are we after in a science of mind? And is consciousness (and the introspection that relies on it) an essential methodological tool for this science, or is it an object of inquiry for this science, or both? If it's a tool, is it a *good* (or even reliable) tool? And what, really, distinguishes first-person and third-person access to the mind? Does a good science of the mind have to use, or make sense of, both? Like his contemporaries, Malebranche does not get worked up, as we do, about any kind of metaphysical problem of consciousness. But he may be one of the first raise a host of serious epistemological and methodological problems about consciousness. And for that he does indeed deserve a place in the history books.

Malebranche's views are so strange, and so theologically driven, that it is easy to write him off as an irrelevant crackpot. It is true that his views about the signature phenomena of the mind (intentionality, representation, perception, idea, and consciousness) are deeply motivated by theological concerns, and in particular the view that the human mind is not a light unto itself but must be enlightened by God. That does indeed make him seem an unlikely source for fruitful inquiry into the mind in an age of reason and science. The result, however, is an (admittedly unusual) account of the mind that shines a fresh light on the conceptual connections among these phenomena, and so, quite possibly, on some of the darker recesses of our own thinking about the mind.[83]

REFERENCES AND SHORT TITLES

Primary Texts

Translations of primary texts are my own but they have been greatly influenced by the standard English translations listed below.

---

[82] *Search* III–II.viii.4, OCM I 451/LO 238.

[83] I am grateful to many people for their helpful comments, suggestions, and challenges as I was writing this paper. Thanks to Martha Bolton, Don Garrett, Sean Greenberg, Jeff McDonough, Kurt Smith, Daniel Sutherland, and to the audiences at Western Ontario, the University of Virginia, and the Mid-Atlantic Conference in Early Modern Philosophy. A special thanks to Michael Della Rocca, Karen Detlefson, Andrew Janiak, Sukjae Lee, and Tad Schmaltz, who suffered through my initial attempt to work through this material, and who provided a great deal of interesting discussion and theatrics.

AT: Tannery, Paul and Adam, Charles, editors. *Oeuvrew de Descartes*. 12 volumes. Paris: Vrin, 1964–1975. Cited by volume and page number.

CC: Malebranche, Nicolas. *Conversations Chrétiennes*. In OCM IV. 3rd edition, 1994.

DM: Malebranche, Nicolas. *Entretiens sur la Métaphysique et sur la Religion.* In OCM XII. 4th edition, 1991. Cited by dialogue and section number.

Elucidations: Malebranche, Nicolas. *Eclaircissements*. In OCM III. 2nd edition, 1976.

JS: Jolley, Nicholas and Scott, David, translators and editors. *Dialogues on Metaphysics and on Religion*. New York: Cambridge University Press, 1997.

LO: Lennon, Thomas and Olscamp, Paul, translators. *The Search After Truth*. Columbus, OH: The Ohio State University Press, 1980.

OCM: Robinet, André, editor. *Oeuvres Complètes*. 20 volumes. Paris: Vrin, 1958–. Cited by volume and page number.

Réponse: Malebranche, Nicolas. *Réponse de l'Auteur de la Recherche de la Verité, au Livre de Mr. Arnauld, des vrayes and des fausses Idées.* In OCM VI–VII. 2nd edition, 1978.

Search: Malebranche, Nicolas. *De la Recherche de la Vérité*. In OCM I–II. Volume I: 3rd edition, 1991. Volume II: 2nd edition, 1974. Cited by book (part, where applicable), chapter, and section number.

Trois Lettres: Malebranche, Nicolas. *Trois Lettres de l"Auteur de la Recherche de la Verité, touchant la Defense de Mr. Anrauld contre la Réponse au Livres des varies et fausses Idées*. In OCM VI–VII. 2nd edition, 1978.

VFI: Arnauld, Antoine. *Des Vrais et Fausses Idées*. Paris: Fayard, 1986.

Spinoza on Teleology, Value, and the Unity of Mind

Charles Jarrett

Introduction

Spinoza seems to think that we are purposive beings in a purposeless world.

God created the world, not in order to achieve anything,[1] but out of necessity.[2] So the world necessarily exists, but it is not "*for* anything." It is thus not "for us." In plain terms, it has no point and no direction. It is "going nowhere."

We, however, are almost always going somewhere. Spinoza grants that we have goals, that is, that we desire things, even if these desires are merely efficient causes. Indeed, he himself sets out an ideal, in relation to which things are good or bad insofar as they promote or hinder our attainment of it.

In this respect, Spinoza seems to be our contemporary, rather than a philosopher of the early modern age. He expresses a contemporary view, at least, because it is now so often thought that the world operates on principles that have no special

[1] See E I App. This paper generally follows the style of abbreviation used in Yovel (1999).

E = *Ethica* (Ethics). "E" is followed by part numbers (I–V) and one of the following:

App = appendix

ax = axiom

c = corollary

cap = caput (heading in E IV App.)

def = definition

def.aff. = definition of affect (in E III)

d = demonstration

exp = explanation

gen.def.aff. = General definition of affect (in E III)

lem = lemma

p = proposition

post = postulate

Pref = preface

s = scholium

Thus "E I p14c2", for example, refers to the second corollary to proposition 14 of the first part of the *Ethics*, and "II p10cs" to the scholium following the corollary of II p10. A comma indicates "and." So "E IV p1,d" refers to proposition 1 and its demonstration, in E IV. See 'Abbreviations' at the end of this paper for other short forms.

[2] E I p16. ("*ex necessitate divinae naturae*." G. II.60, 18.)

J. Miller (ed.), *Topics in Early Modern Philosophy of Mind*, Studies in the History of Philosophy of Mind 9, 131–149.

concern for us. We exist in a world that has for a while been conducive to, or at least compatible with, our continued existence. But we might just as easily follow the dinosaur and perhaps we should even expect it. For as science fiction scenarios often suggest, our extinction would not require much. A meteor's collision with the earth, a massive volcanic eruption, or extensive global pollution might be enough.

If this is Spinoza's news, then we must apparently conceive of ourselves as in some measure alienated from "the world" and as operating on general principles that do not apply to it. The "modern" thought is that we are beings who constantly strive for something and are always looking toward the future or "projecting." "The world" itself, however, does not do this. It just is. So each of us is a *for itself* in the midst of the *in itself* and not "at home" in the world.

A fundamental question about Spinoza's philosophy is whether this reading can be sustained. For Spinoza's rejection of final causes may be thought to entail that we are not purposive beings at all, or at least not more purposive than a rock.[3] His conception of ethics as a theory of rationality and of us as rational beings, in contrast, seems to portray us as quite unlike other parts of nature.

A fundamental question of philosophy is whether we can succeed in our attempt to understand ourselves as parts of nature, explicable on just the same general principles as all other things. How, within such a framework, are ethics and rationality to be conceived? At issue is the extent to which human beings transcend the rest of nature.

In this paper, I hope to contribute to a solution by considering Spinoza's conception of the unity and nature of the human mind. My aim is to characterize and assess Spinoza's conception of this, as well as of the roles that teleology and value play in it. Section 1 is concerned with Spinoza's views on the unity and nature of the mind, while Sections 2 and 3 deal, respectively, with issues about teleology and value. Section 4, finally, considers a threat to Spinoza's account of the unity of the mind that arises from his apparent transcendence of ethics, and of the temporal world, in the last portion of Part V of the *Ethics*.

1 The Unity of the Mind

The general question, "What accounts for the unity of a thing?" or "In virtue of what is a thing a unity?", is not completely transparent. On the face of it, what makes a thing one mind or one coin, for example, is whatever it is that makes it a mind or a coin. If we can set out the conditions under which something is a coin, or what it is to be a coin, then we will have set out the conditions under which something is one coin. More generally, it seems, what makes a thing one, or a

[3] For a comparison of us with a rock, see Ep 58, S 909. For additional remarks on this, see Carriero (2005, pp. 135–136).

unity, is whatever it is that makes it an F, where "F" is a count noun, such as "coin" and "horse," as opposed to a mass term such as "money" and "glue." Being an F is then what makes something one F.[4]

Spinoza seems to agree with this account of the unity of a thing, as indicated in his early CM and, more explicitly, in Ep 50.

In CM I, vi Spinoza considers "the one, the true and the good." Spinoza's opponents, that is, "almost all metaphysicians," maintain that "unity signifies something real outside the intellect."[5] His diagnosis is that they "are confusing beings of reason with real being" and he expresses his own view as follows:

> … [U]nity is in no way distinct from the thing itself or additional to being and is merely a mode of thinking whereby we separate a thing from other things that are similar to it or agree with it in some respect.[6]

Its opposite, plurality, is also merely a "mode of thinking."

Spinoza also notes here that we say God is one and unique, but he adds, "In truth, … we might perhaps show that God is only improperly called one and unique …."[7]

In Ep 50, written in 1675 to Jarig Jelles, he explains this last remark. He says there that we call a thing "one" or single with respect to its existence, not its essence, and he adds:

> For we do not conceive things under the category of numbers unless they are included in a common class. For example, he who holds in his hand a penny and a dollar will not think of the number two unless he can apply a common name to the penny and dollar, that is, pieces of money or coins.[8]

Thus, since "we can form no universal idea" of God's essence, we cannot properly say that God is one or unique.[9]

To determine what constitutes the identity of a mind, we must then know, most fundamentally, what kind of thing a mind is.

A human mind, according to Spinoza, is God's idea of the human body,[10] and so a human mind is one, or is a unity, because it is one idea. It is the mind that it is, and is distinct from other minds, because of the fact that it is the idea of one

---

[4] A distinct, but related question concerns the distinction between one thing that is F and another, that is, between one unity and another of the same kind.

[5] S 186.

[6] S 186.

[7] CM I vi, S 187. ("*At verò si rem accuratiùs examinare vellemus, possemus forte ostere Deum non nisi impropriè unum, et unicum vocari, sed res non est tanti, imò nullius momenti iis, qui de rebus, non verò de nominibus sunt solliciti*." G I.246, 9–13.)

[8] Ep 50, S 892.

[9] *Ibid*. He writes: "Now since the existence of God is his very essence, and since we can form no universal idea of his essence, it is certain that he who calls God one or single has no true idea of God or is speaking of him very improperly."

[10] E II p13.

body rather than another. Minds can thus be distinguished from each other by distinguishing their bodies.[11]

Spinoza's thesis is quite general, however. He holds that God has an idea, not just of each human body, but of each body, and so all things are animate, "although in different degrees."[12] He also maintains that the idea of each human body is composed of ideas of the parts of that body,[13] and so the human mind is composed of many minds. It also contains ideas of everything that happens in the human body.[14]

But in what does the unity of a human body consist? In the physics of E II, set out between II p13 and II p14, a structural notion is used. It consists in the preservation of the ratio of movement among the parts. He writes:

> When a number of bodies … form close contact with one another through the pressure of other bodies upon them, or if they are moving at the same or different rates of speed so as to preserve an unvarying relation of movement among themselves, these bodies are said to be united with one another and all together to form one body or individual thing, which is distinguished from other things through this union of bodies.[15]

In both the *Ethics* and the KV, he also describes this as a relation or ratio (*ratio*) of motion and rest among a body's parts.[16] In the latter work he illustrates it as a ratio of, for example, 1 to 3.[17]

Della Rocca provides a useful example of this. He contrasts a complex individual, and Spinozistic unity, with a mere collection by comparing his dining room chair with the collection consisting of his chair and his telephone.[18] The parts composing his chair "have a tendency to stay in a certain overall relation,"[19] while the parts of the "chair–telephone" do not. If we move the back of the chair, for example, the other parts move as well, but if we move the telephone, nothing, or almost nothing, happens to the chair. Carriero calls these unitary things "pattern-like" and suggests that we think of "watches, ferns, snails, and even the human body, in Descartes' and Spinoza's metaphysics, along the lines of especially complex and stable tornadoes."[20]

---

[11] It should be noted that this holds only for minds of things that are regarded as bodies, which are conceived under the attribute of extension. When a thing is conceived under another attribute, its mind is distinct from God's idea of the body. For each thing has an infinity of minds, according to Ep 66.

[12] E II p13s.

[13] E II p15,d.

[14] E II p12.

[15] E II Definition (just prior to Axiom 3, between lemmas 3 and 4), S 253.

[16] See E II lem 5, S 254 and KV II Pref, note 1, S 60–61.

[17] KV II Pref, note 1 (Sections 12 and 14), S 61.

[18] See Della Rocca (1996, pp. 207–209).

[19] *Ibid*., p. 207.

[20] Carriero (2005, p. 123).

Spinoza's account of a mind thus looks reductive. The identity or unity of a mind is accounted for by saying that it is the idea of one body and it is one body because its parts bear certain structural and causal relations to each other. So the question of the unity of the mind is ultimately answered by appeal to purely physical and, on the face of it, "mechanical" principles.[21]

This reductivist allegation, however, ignores the fact that Spinoza uses a notion of representation, and of the content of an idea, since a mind is an idea, in God, *of* the body. Indeed, it is God's affirmation of the existence of the body and this affirmation endures in time just as long as the body does. On the other hand, Spinoza explicitly maintains that the mind and the body are "one and the same thing, although conceived in different ways" and that "the order and connection of ideas is the same as the order and connection of things."[22] The latter by itself seems to entail that no principles govern the structure of the mind or the sequence of its ideas except "mechanical ones" and that the mind is in effect a mental machine.

Despite this, Spinoza might be thought to import a notion of teleological unity into his concept of the mind, as well as of the body, when he characterizes them, and every individual, as having one most basic conatus or desire. Very much at issue, however, is whether his concept of a conatus is a "purely mechanical" notion, as opposed to a teleological one.

The differences between these types of notions are not entirely clear, but an illustration may be useful. We ordinarily think of a rock, for example, as not having teleological unity. It is not trying to do anything, nor is it "for" anything. It is a unity, one rock rather than a mere heap or collection, because its parts stay together. You can pick it up and throw it and all of the parts go with it. But the parts are merely causally connected to each other.

In contrast is a teleological unity. We think of animals, for example, as characteristically doing things, and having parts, that promote their survival and reproduction. They seem not merely to produce, but to aim at, things in the future. Another type of example is an artifact, such as a home heating system, a story, a grocery list, or a corporation. A grocery list, for example, may be one list, even if it starts on one sheet of paper, but is continued on another or on a chalkboard. A corporation may be a unity, even if it is physically scattered in different branch offices with poor communication between them. The unity of such entities is surely inexplicable solely in mechanical terms (that is, in terms of the concept of a "blind" physical cause).

Although Spinoza characterizes the essence of a body (and mind) as a conatus or endeavour to persevere in existence, in E III p6 and p7, this does not by

[21] The notion of a "mechanical" principle is by no means transparent. It is perhaps tempting to define it as any explanatory principle that relies solely on the concept of an efficient cause, but this is inadequate for several reasons. More promising would be the idea that a mechanical principle is one that encompasses principles that account for states of things and changes in them without reliance on the concept of a final cause.

[22] E II p7, S 247.

itself settle any issue about teleology. For he ascribes this conatus to the uniform rectilinear motion of a body as well. In doing so, he seems to regard such motion as a state of the body, which requires the citation of no separate cause to explain its continuation. In this respect it is unlike an event, such as a change in the direction or speed of the motion, which does require a separate cause.

Another notion of the unity of the mind, or rather of a union of the mind with something, is expressed by Spinoza in his early *Korte Verhandeling*. In the KV Spinoza holds that we are reborn when we replace our love of the body with the love of God and he speaks here of a union, first with one, then the other. He maintains that if we love the body, and are united with it because of that, then when the body dies, so does the soul or person. If we love God, however, then we are united with him, and continue to exist in time after the body is destroyed.[23]

Spinoza here seems to express the thesis that you are what you love, or are defined by what you most love, and this can change. You "become a different person," we might informally say. But what does this mean? On one account, you are initially one individual who is primarily concerned with one thing and then you, the same individual, become primarily concerned with another. You become a different kind of person. On another account, you would not survive the change. It would mean, it seems, that you actually cease to exist and a new individual arises.

The same issue arises in the case of the change from infancy to adulthood and the change that some adults undergo when they lose their faculties. Spinoza considers both cases and we will turn to them in Section 3.

Spinoza's account in the KV was abandoned, however, in the *Ethics*. There he holds that the mind cannot continue to exist in time after the destruction of the body. Despite this, he retains a conception of the immortality, or rather of the eternity, of at least the human intellect. Further discussion of this is reserved for Section 4.

2 Teleology

It is tempting to think that the unity of the human mind must consist in its having one primary function or one primary goal. If we think of ourselves as purely biological beings, for example, we may regard perception, emotion, thought, and even consciousness itself, as existing mainly to promote our own continued existence or successful reproduction. If we instead think of ourselves as psychological subjects, we may regard the great diversity of our interests, desires, and activities as unified only by having a single dominant goal, to which all other

[23] See KV II.Pref, note 1, Section 15, S 61 and KV II.xxiii, S 95. Also see KV II.xix, S 89, where the union of the mind with God is regarded as being released from the body; and KV II.xxvi, S 101–102, where human freedom is regarded as independence from external causes.

interests are, or should be, subordinate. If, finally, human minds or human beings are essentially intellects, or rational beings, we may regard ourselves as having one primary function or goal: to understand more.

Spinoza is an insistent advocate of this last view. He holds that the human mind, at least, is essentially rational and that its only real power is the power to understand or to have knowledge.[24] In addition, he sets out a single ideal, in relation to which any assessment of things as good or bad must be made.[25] This ideal or goal is the attainment of knowledge or understanding.[26] Spinoza thus posits a single primary goal as the basis of ethics and he takes the human mind to have a unitary power.

Despite this, Spinoza seems to eschew "teleological" conceptions of the unity of the human mind in favour of a "purely causal" account. He seems, that is, to endeavour to reduce final causes to efficient causes, and so to reduce teleological explanations to what we call "causal" explanations (explanations in terms of efficient causes).

The question whether Spinoza fully succeeds in dispensing with final causes, or in reducing them to efficient causes, is a difficult one and his conception of teleology has been the subject of much recent work.[27] A large part of the problem consists in clarifying what Spinoza takes a final cause to be, how he uses concepts such as that of desire, and what, finally, we take a teleological conception or explanation to be.

Spinoza's views on final causes are expressed in many places. Perhaps the most important are found in the remarkable appendix to Part I, the Preface to E IV, and, quite briefly, in E IV def7.

Spinoza is not an eliminitivist with respect to all final causes, as E IV Pref and IV def7 indicate. He is rather a reductivist. IV def7 states: "[B]y the end for the sake of which we do something, I mean appetite."[28] E IV Pref maintains: "What is termed a 'final cause' is nothing but human appetite insofar as it is considered as the starting point or primary cause of some thing."[29]

---

[24] See, for example, E IV p26, S 333 ("Whatever we endeavour according to reason is nothing else but understanding and the mind, insofar as it exercises reason, judges nothing else to be to its advantage except what conduces to understanding.") and E V p20s, S 373 ("… the power of the mind is defined solely by knowledge …").

[25] See, for example, E IV Pref, S 322: "So in what follows I shall mean by 'good' that which we certainly know to be the means for our approaching nearer to the model of human nature that we set before ourselves, and by 'bad' that which we certainly know prevents us from reproducing the said model."

[26] See, e.g., E IV p 26, S 333 and his restatement of it in E IV p27,d, S 334: "The mind, insofar as it exercises reason, seeks nothing else but to understand …."

[27] See, for example, Bennett (1984, pp. 213–230; 1990, pp. 49–52; 2001, pp. 207–223), Curley (1990), Garrett (1991), Della Rocca (1996, esp. pp. 252–257), Jarrett (1999), and Carriero (2005).

[28] S 323.

[29] S 321.

Carriero has very helpfully clarified Spinoza's rejection of final causes and I, for one, am indebted to his work on this.[30] He does this in part by setting out and elucidating Spinoza's primary target, namely an Aristotelian conception, as expressed by Aquinas.

Several closely connected elements of the Aristotelian and Thomistic conception are especially important, as Carriero informs us. For example, final causes or ends account for, or make possible, the causality of the efficient cause,[31] and efficient causes are not "blind."[32] In addition, efficient causes have a "natural end."[33] They do not just go on forever; instead, they cease to operate when the natural end (the actualization of something in potentiality) is attained.

Spinoza's rejection of the last of these can be illustrated and supported by noting that he conceives of the conatus and essence of a finite mode as an endeavour to persevere in existence *indefinitely*. It is also confirmed by an especially important example that Spinoza provides concerning the extinction of desire.

It is tempting to think that at least some desires or "final causes" are naturally extinguished when their ends are attained. It seems to be built into the concept of a desire that it ceases to exist when it as "satisfied" (or thought to be satisfied). So if you want to do something, then when you have done it, or believe you have done it, the desire to do it "automatically" ceases. Your thirst is "quenched," we say, and so you no longer desire to drink.

This seems to be an element of our ordinary (teleological) concept of desire, at least in many cases. So desires are precisely the sort of thing that bring about their own destruction.

An important feature of Spinoza's thought about this is that he does not regard a desire as automatically extinguished by attainment of the goal. He notes in E III p59s that when we enjoy something that we love or seek, the body's condition changes and we begin to want other things. He says:

> For example, when we think of something that is wont to delight us with its taste, we desire to enjoy it, to eat it. But while we are thus enjoying it the stomach is being filled and the body is changing its condition. If therefore, with the body now in a different condition, the image of the said food is fostered by its being set before us, and consequently also the conatus or desire to eat the food, this conatus, or desire, will be opposed by the new condition of the body, and consequently the presence of the food which we used to want will be hateful, and this is what we call Satiety (*fastidium*) and Weariness (*taedium*).[34]

What Spinoza here gives is a "cybernetic" account. He regards a desire as destroyed, not merely because it is "satisfied," or because the end is achieved, but

[30] Carriero (2005).

[31] Carriero (2005, p. 113).

[32] *Ibid*., p. 116.

[33] *Ibid*., p. 109.

[34] S 310–311.

because it causes some change that itself turns back to affect, and indeed destroy, the desire. We have a "causal feedback loop," like that found in a home heating system.

A further comparison of Spinoza's views with current ideas about teleological explanation, even if not precisely formulated,[35] may also be useful. Such explanations are comprised of two types: (1) functional, where we say what good it does, *for* something; (2) intentional, where we say what good it is thought by some agent to do.

Functional explanations answer the question, "What is it for?" and this practically means, "What good does it do?" A thing that is explained in this way then has a productive role or a contribution to make, as when we say that the function of the eyes is to enable us to see. Such explanations seem to proceed primarily by describing or redescribing the thing in terms of its effects.

Functional explanations, however, do not seem completely reducible to "mechanical explanations," or explanations in terms of efficient causes. For not just any effect of a thing can be cited in a statement of its function. In the case of living things the effects are restricted to those that promote continued life or reproduction. In the case of artifacts, they are typically the things we want them to do.

When we ask the question, "Why did he die?", we can understand this in two ways and we can get two types of answer. On one, we are asking for the cause of death, as set out, for example, on a death certificate. On the other, we are asking for the point of it or what good it does. We seek to know why it is somehow fitting or even what should be, even if we are ignorant of the details.

Spinoza himself considers this type of example in E I App. He died because the stone fell off a roof and hit him. The stone fell because of a strong wind and it hit him because he was passing that way—and so on. Only efficient causes are employed.

That it is the will of God, Spinoza considers and rejects. God's will, as ordinarily conceived, is the asylum of ignorance. His own position is that the will of God is his understanding, which is necessitated for no end. So Spinoza here seems simply to reject a functional explanation. It serves no good and has no purpose; it just happens and is necessitated by preceding causes.

A more promising field for functional explanation is found in biology. Consider the question, "Why do we have eyes?" One answer is, "so that we can see" or "in order to see." A more complete answer might also be provided. It is so that we can see dangers and avoid them, get our dinner, etc.

Spinoza's answer, on the face of it, is that having eyes enables us to see and avoid dangers, but we have them because of the causal processes that lead to their existence. We should turn to embryology, not theology, to answer the question, for seeing (in the case of any individual) is the effect of having eyes, not its cause.

[35] Carriero (2005) characterizes current thought of final causes, not unjustly, I think, as "a rather blunt instrument of contemporary and somewhat free floating notions of a teleological or functional explanation …." (p. 106)

Functional explanations of this sort also import the concept of a defect and so Spinoza seems committed to their rejection.[36]

Intentional explanations, in contrast, are typically used to explain human action by citing the desires and beliefs of the agent.

Spinoza discusses this in E IV Pref, where he speaks of residential construction. He writes:

> What is termed a "final cause" is nothing but an appetite insofar as it is considered as the starting point or primary cause of some thing. For example, we surely mean no more than this, that a man, from thinking of the advantages of domestic life, had an urge to build a house. Therefore, the need for a habitation insofar as it is considered as a final cause is nothing but this particular urge, which is in reality an efficient cause ….[37]

Here we have a conception of a human desire as a cause and an apparent attempt to "reduce" intentions or desires to purely causal factors. It differs somewhat from Davidson's well-known account[38] because Spinoza takes imagining something, or having ideas, as causes of the desire, which in turn causes the act. Davidson, in contrast, apparently regards desires and beliefs as joint causes of the act.

It looks as if Spinoza is committed to a reductionistic account of intentional explanations to purely causal explanations and to an outright rejection of functional explanations, at least insofar as the end—a future or abstract state—is cited to explain some biological feature.

3 Relative and Absolute Value

Spinoza often insists that things are good or bad only "relatively" or "in relation to" something and he in fact goes so far as to apply this doctrine to God.[39] In E IV Pref, for example, he maintains that he will use the terms for what helps or hinders our becoming like the ideal that we have set out.

To what extent is Spinoza's conception of the relativity of value tied to a denial of purposiveness? Indeed, how are teleology and valued connected? Some light can be shed on this, perhaps, by considering the project that Paul Grice sets out in the last of his three Carus Lectures.[40]

---

[36] See Carriero (2006, pp. 126–131) for helpful remarks about the notion of a fault in both the scholastic Aristotelian tradition and in Spinoza.

[37] S 321.

[38] See Davidson (1963).

[39] See KV I.vii and CM I.6 on God's goodness. See Jarrett (2002) regarding the variety of types of relativity that Spinoza employs and see Miller (2005) for more on Spinoza's views concerning the relativity of value.

[40] Grice (1991, pp. 23–91). The first two of Grice's Carus Lectures raise problems for the theses that (i) there is no objective and no absolute value and (ii) there are no categorical imperatives, as set out by J.L. Mackie and Philippa Foot. See Mackie (1977) and Foot (1978). Both of these theses seem to be endorsed by Spinoza.

A useful outline of central parts of Grice's views is provided by Judith Baker and a mere repetition of it, along with her lettering, is as follows.

> (a) Grice supposes that living things are described in terms of final causes or "finality features," that is, they are characterized in terms of what they do, and what "they aspire to."[41]
> (b) Living things and their behavior can be evaluated in terms of "finality or function." This includes an evaluation of them as being good of a kind.[42]
> (c) Human beings have rationality accidentally. This is a "capacity and concern" that attitudes, beliefs, etc. be justified and a concern that some be "absolutely validated or of absolute value."[43]
> (d) Some human beings, via "metaphysical transubstantiation," become persons, a "new kind of creature" that is essentially rational.[44]
> (e) A good person is then "one who is good at her function or metier of finding reasons."[45]
> (f) We "project" our "thinking-valuable" on things and value is legitimately or justifiably transmitted to these things by a good person.[46]
> (g) Some of this transmitted value is absolute. A person who is good *qua* person may then be a main example of something that is of absolute value.[47]

It is clear that Spinoza balks, perhaps solely, at items (g) and (c). Spinoza thinks that all value is relative and he takes a good person to be good either as a means to an increase in our own understanding or as someone who matches our idea of the type (or both).

Item (a) seems acceptable to him, at least nominally, since (for example) he takes a human mind to aspire to greater understanding; but much, clearly, turns on how a "finality feature" is understood. (Spinoza, however, evidently diverges from Grice in thinking that everything is alive.) No doubt (b) is acceptable, since Spinoza evaluates some people as good or bad (i.e., strong or weak). (e) and (f) seem unproblematic as well. Finally, item (d) is quite problematic, but I will suggest below that Spinoza may well have accepted it.

Part of Grice's idea seems to be that rationality may have originated and been sustained by what we think of, at least initially, as "mechanical" principles. A genetic mutation, or a series of them, might have arisen and been passed down via purely physical mechanisms. Thus a certain biological species, *homo sapiens* (and maybe others), has arisen that has intelligence, or has greater intelligence than most others.

---

[41] Grice (1991, p. 5).
[42] *Ibid*., pp. 5–6.
[43] *Ibid*., p. 6.
[44] *Ibid*.
[45] *Ibid*.
[46] *Ibid*.
[47] *Ibid*.

The idea is that rationality exists, initially, as a biological advantage, but that, having arisen, it cannot be confined to a fixed end, such as survival or reproduction; instead, it imports, or makes possible, questions such as what to take as an ultimate goal. As Grice puts it, the genitor "gets more than he bargained for."[48] This naturalistic story ends, however, when metaphysical transubstantiation occurs, as set out in Baker's item (d), for then a person arises who is essentially rational. This is a non-biological type.[49] According to Grice, there is then a guarantee that, if further conditions are met, there is absolute value and finality for which there is no "mechanistic substitutability."[50]

What Grice thus supposes or suggests is that the postulation of absolute value is tantamount to the thesis that there is teleology that is not explicable solely in terms of efficient causes.[51] It is not, as he puts it, "mechanistically substitutable."[52]

In addition, he introduces the interesting idea that an individual may possess rationality non-essentially, as a member of the biological category *homo sapiens*, but that he or she can later come to possess it essentially. In such a case, an individual human being would "turn into" a rational being, a person, and in the process the human being would cease to exist. Indeed, his view, as he puts it, "would allow a thing x and a thing y to be identical at a certain time but to be not identical at a different time, when indeed one of the things may have ceased to exist."[53]

This is an example of what he calls "metaphysical transubstantiation," and it involves the construction and use of a concept of relative identity. We may then say, for example, that I was once a certain infant, or even a certain member of a biological type, but I am no longer. Although I once was the same as that being, I am now not.

Now this program is of interest, and of use, in an attempt to understand Spinoza for a variety of reasons.

(1) At one stage, Spinoza supposes, or gives the impression that he supposes, that the essence of a human being is the endeavour or desire to persevere in existence. That is, human beings, like all beings except God (perhaps the sole exception), have this as their primary or even highest goal. This may express or rely on a conception of human beings as a biological category, but since all things are "animata," it is not clear that any important contrast remains, on this point, between human beings and any other modes.

---

[48] Grice (1991, p. 85).

[49] *Ibid*., p. 84.

[50] *Ibid*., pp. 89–91.

[51] *Ibid*., pp. 78–79.

[52] *Ibid*., p. 88.

[53] *Ibid*., pp. 81–82.

There is also, however, a stage or a point at which Spinoza passes beyond mere biological survival and supposes that we are essentially rational beings. He expresses this conception of us in the TP V, 5, for example, as follows.

> So when we say that the best state is one where men pass their lives in harmony, I am speaking of human life, which is characterised not just by the circulation of the blood and other features common to all animals, but especially by reason, the true virtue and life of the mind.[54]

The question or problem that this raises is how our goal of survival as biological beings, or our perseverance in mere existence, common to at least all finite modes, is related to our highest goal, and good, as rational beings. Spinoza's answer is found in E IV p26d, where he maintains that

> this endeavour of the mind, by which the mind, insofar as it reasons, endeavours to preserve its own being is nothing other than to understand. …[55]

Spinoza thus holds that it is our survival *as* rational beings that we most want, but he muddies the waters, at least, when he supposes that death is not so important. He holds, for example, that the free man scarcely thinks of death[56] and that those with a mind the greatest part of which is eternal scarcely fear death.[57] Indeed, he maintains that a rational being will in fact sometimes act in ways that are known will lead to his or her own death. Such a being will not lie, for example, even to save his or her own life.[58]

There is thus some tension, at least, between Spinoza's initial supposition that our conatus is an endeavour or tendency to remain in existence, or to remain alive, and his later proclamation that our essence consists in rationality.

It might well be that Spinoza thinks that a transition occurs. Perhaps it is even quite evident that he thinks this, because he supposes that we are barely conscious and, it seems, not fully rational beings at birth. He holds that infants and children, like "the ignorant," are almost entirely unconscious of themselves, things, and God, while the wise are very much conscious of them.[59]

In E V p39s he speaks as if the transition to adulthood involves the genesis of a new body and hence, we must suppose, a new mind. He writes:

> In this life, therefore, we especially endeavour, that the body of infancy, insofar as its nature allows and is conducive to it, be changed into another, that is capable of many

---

[54] S 699.

[55] "… *hic Mentis conatus, quo Mens, quatenus ratiocinatur, suum esse conatur conservare, nihil aliud est, quàm intelligere* …." (G II.227, 16–18, my translation above) Spinoza's argument for this has been severely criticized in Bennett (1984, p. 305).

[56] E IV p67,d.

[57] E V p39s.

[58] E IV p72s.

[59] See E V p39s, V p42s.

> things, and that it be related to a mind which is very conscious of itself, and of God, and of many things ….[60]

The opposite can also apparently occur, as indicated by the case of the Spanish poet who lost his memory. In IV p39s, where he discusses this, Spinoza says:

> For I do not venture to deny that the human body, while retaining blood circulation and whatever else is regarded as essential to life, can nevertheless assume another nature quite different from its own.[61]

Indeed, he says, experience seems to teach that an adult's body may die, or cease to exist, "without turning into a corpse,"[62] as indicated by the case of the Spanish poet.

Thus Spinoza seems to think that the transition consists in the generation of a new individual, rather than a change in a continuously existing one, that is, that it is not the same individual who strives at birth to live, but only later strives to understand, or to live as a rational being. His view seems rather to be that the earlier body dies and a different body begins to exist.

But perhaps nothing prevents Spinoza from taking a Gricean line here and holding that the now rational individual was the same as the infant, but now is not. Indeed, this would obviate the problem that in advocating the transition from ignorance or infancy to wisdom, in IV p39s and V p39s, Spinoza advocates our own death.[63]

(2) Another strand in how this helps us to understand Spinoza is found in the mere fact that Spinoza rejects any notion of absolute value as well as all "final causes," except insofar as they are efficient causes. There is even a passage, which is perhaps not entirely clear, where Spinoza seems to equate the postulation of a

---

[60] "*In hâc vitâ igitur apprimè conamur, ut Corpus infantiae in aliud, quantùm ejus natura patitur, eíque conducit, mutetur, quod ad plurima aptum sit, quodqúe ad Mentem referatur, quae sui, & Dei, & rerum plurimùm sit conscia ….*" (G II.305, 28–32, my translation above) See also KV II Pref (footnote 1, subsection 10): "This body of ours, however, had a different proportion of motion and rest when it was an unborn embryo; and in due course, when we are dead, it will have a different proportion again; nonetheless there was at that time [before our birth], and there will be then [after death] an idea, knowledge, etc. of our body in the thinking thing, just as there is now, but by no means the same [idea, etc.], since it is now differently proportioned as regards motion and rest." (S 60–61)

[61] S 342.

[62] S 342.

[63] In his 1991, pp. 81–82, Grice calls this "the Grice–Myro theory of identity." It is developed by George Myro in his 1985. See Myro (1986). See Gallois (1998) for further development of a notion of identity relative to a time and a consideration of objections. A conception of relative identity, although not relative to a time, might also be of use in resolving this problem as well as the problem of the attributes. See Geach (1972, pp. 238–249; 1973) for more on this notion of relative identity. For a sample of criticisms, see Wiggins (2001), and Bennett and Alston (1984).

"metaphysical good" with the thesis that the conatus of a thing is "separate from" the thing itself.[64]

What Spinoza may mean here is that if the conatus is separate from the thing, then there is a goal or "final cause" that is not merely an efficient cause. There is then something antecedent to and independent of the conatus (a "detached" final cause) that both accounts for the conatus and is regarded as "metaphysically" or absolutely good.

Absolute or non-relative goodness would then exist and it would explain our desire for it as well as our action toward it, thus "reversing" the order of nature. Contrary to Spinoza's own dictum,[65] we would desire something because we judge it to be good, rather than the reverse. Indeed, we would desire it because it is good.

Grice accepts Spinoza's connection between absolute value and "detached finality," but he jumps the other way. Grice holds that absolute value arises, or can justifiably be regarded as arising, just when there are "finality features" that are not "mechanistically substitutable," that is, when there is an area where there are final causes or goal directedness that cannot be fully explicated or accounted for in terms of cybernetic or mechanistic notions.

So both seem to tie the acceptability of final causes that are not themselves merely efficient causes to the acceptability of absolute value and also, I think, to the acceptability of categorical imperatives.

4 Metaphysics and Ethics: Kantian Troubles

A final difficulty must be faced, however, if we are to ascribe to Spinoza a coherent conception of the unity of the human mind.

It is perhaps commonly and naturally thought that what is of most significance in Spinoza's *Ethics* is his attempt to provide a rational guide or basis for a "way of being in the world." This may also be characterized as an attempt to say what attitudes and values we will have, and how we will act, insofar as we are rational.

At least as important, however, and certainly essential for the purpose of obtaining a relatively clear conception of Spinoza's philosophy as a whole, is the fact that the *Ethics* does not come to an end at the point (in V p20s) at which Spinoza has finished speaking of "the present life." Spinoza proceeds instead to characterize what he calls "freedom of mind or blessedness," and in so doing (from V p21 to p40) he takes or attempts to take the stance of one who is, as it might well be put, in 'another world.'

---

[64] CM I vi, S 188: "[Why some have maintained that there is a metaphysical good.] But those who keep seeking some metaphysical good not qualified by any relation are laboring under a misapprehension.... For they are making a distinction between the thing itself and the conatus [striving] to preserve its own being, which every thing possesses, although they do not know what they mean by conatus."

[65] E III p9s, S 284.

There is, as usual, no general agreement among commentators concerning the precise nature of Spinoza's distinction between the present life and the life of blessedness, nor is it my intention here to settle this by detailed examination of the text.

What I wish to do instead is to explore this dual tendency in Spinoza's thought, employ one interpretation of it in an attempt to elucidate Spinoza's metaphysical foundations of morals, and then merely note, finally, that Spinoza encounters a Kantian-like problem in retaining a unitary concept of the human mind.

The central metaphysical and epistemological basis or, if you prefer, expression of Spinoza's two standpoints is found in the distinction Spinoza explicitly sets out in V p29s. There we find that "things are conceived by us as actual in two ways[:]"[66] they are conceived "to exist in relation to a fixed time and place"[67] and they are conceived "to be contained in God and to follow from the necessity of the divine nature."[68] I will occasionally say (as does Spinoza) that in the first way we conceive a thing *sub duratione*, and in the second, *sub aeternitatis specie*. And while I will not argue the matter here, conceiving a thing *sub aeternitatis specie* is conceiving of it as *atemporally*, rather than omnitemporally, existent.[69]

Thus Spinoza may be said to have advocated a theoretical conception or standpoint that 'transcends' our ordinary conception of the world as inherently spatio-temporal. His distinction between the two ways of conceiving things is then not just incidentally related to Kant's distinction between the (positive) concept of a noumenon and the concept of a phenomenon (a point to which we will shortly return).[70]

Spinoza's two standpoints can be more fully expressed as follows. From the first standpoint, each thing endeavours, and has a varying degree of power, to preserve itself and in fact this endeavour is the foundation of virtue.[71] What you endeavour to preserve is yourself, regarded as an idea of an actually existing body in time, that is, as a mind whose sole power is the power to think and form adequate ideas. From this point of view Spinoza characterizes as good (for you) anything that promotes the integrity of your body, and anything that increases the power of your body to persist through a variety of changes and to produce a variety of changes. For the latter increases (or rather is necessarily conjoined with an increase in) your capacity to think or understand. Preserving your body in time and engaging in various activities (such as eating a variety of foods and helping rather than harming other people) are, in short, advocated as extrinsic or instrumental goods. They are good only because and insofar as they contribute to

[66] G. II.298,30, my translation.

[67] S 376.

[68] S 376.

[69] See Jarrett (1990) for more on this and on the relations of Spinoza's views to Kant's.

[70] See Kant (1984), especially "On the Ground of the Division of all Objects into Phenomena and Noumena" (Bk. II, Ch. III), and Kant (1993, esp. Sections 30, 32–34, 45, 57, 59).

[71] E IV p18s; G II.222, 26–27, S 330.

understanding, the "life of the mind," of ourselves as conceived to exist *sub duratione*.

The second standpoint is that from which things are conceived *sub aeternitatis specie* by the intellect, that is, the "better part" of us.[72] This is a standpoint from which there is no memory, imagination, or sense-perception, and no ability to think of objects in relation to (or as in) time and place. There is adequate understanding of the essence of God and of the things that eternally follow from his essence and there is love of God that 'arises' (atemporally) from this knowledge and which, like that knowledge itself, is eternal. Since this is a standpoint that includes no inadequate ideas, it is also a standpoint from which good and evil are inconceivable.[73]

We might add that Spinoza not only characterizes and himself takes these two standpoints (in diverse sections of the *Ethics*), but also has two attitudes both toward himself and toward the world.

On the one hand, his own death as a being in time is to be resisted (rationally) by him as much as possible, and the thought of his own death will certainly be accompanied by the most extreme sadness (which itself is bad). On the other hand, the destruction of himself as a being in time is really unimportant (by V p38s) and, speaking strictly, from the second standpoint, inconceivable (by V p21, 29, and their demonstrations).

Toward 'the world' Spinoza exhibits (as I suppose everyone does) a dual attitude, but it is perhaps preponderantly negative. We can say, I think, that Spinoza noticed and emphasized the 'paradox' of being in the world that consists in this fact: that the things that are most important to us in 'the present life,' that is, those that we most highly value and especially the people whom we most love, are the source of our greatest despair and sadness. The more you love people the more you will despair at any harm that comes to them, and at your loss of them, and the more you will hate and attempt to destroy those who harm them. Love of perishable things, including people, is then a risky position, and Spinoza in fact advocated 'freeing yourself' from such love as much as possible. He writes:

> … [E]motional distress and unhappiness have their origin especially in excessive love toward a thing subject to considerable instability, a thing which we can never completely possess. For nobody is disturbed or anxious about any thing unless he loves it, nor do wrongs, suspicions, enmities, etc. arise except from love toward things which nobody can truly possess.[74]

Spinoza recommends the replacement of such love with the love of something incapable of being destroyed except with the body itself, namely love of God, and this love, as noted in V p20s, "cannot be defiled by any of the faults that are to be found in the common sort of love …."[75] He thus recommends not merely the

[72] E IV App, cap2.

[73] See E IV p65 and p68,d,s.

[74] E V p20s, S 373. Cf. TdIE sec. 9 and 10.

[75] S 373.

destruction of the evil affects of hatred, anger, jealousy, guilt, and so on, but also the destruction of love of transient things—that is, our principal sources (or forms) of 'attachment' to the world. For such love is the main source of those evil affects.

As noted earlier, there is a marked similarity between Spinoza's "two standpoints," based on the distinction between two ways of conceiving things, and Kant's distinction between noumena (in the positive sense) and phenomena. The latter distinction may be briefly described as the distinction between a thing as it is in itself, independently of how it affects us (in space–time), and a thing, insofar as it is conceived as spatio-temporal or merely temporal. It is in this respect that the central similarity, if not identity, of the distinction as found in Spinoza and Kant is to be found. Kant of course maintains that the positive notion of a noumenon, or positive employment of the concept, is illegitimate, at least partly because it leads to antinomies, and that in this sense, the concept of a noumenon is not even known to be the concept of a possible object, or of a thing that could exist. It is only in its negative sense, or in use of the concept solely as a limiting concept, that the notion is legitimate and, indeed, indispensable. It is in this respect, as also in the allied Kantian doctrine that there could be no "intellectual intuition," that Spinoza's position diverges radically from Kant's.

Spinoza, then, in taking two standpoints seems to regard us as two radically different kinds of beings. The standpoint from which things are conceived *sub aeternitatis specie* is a standpoint that is both within and concerned with the noumenal world. In the noumenal world there are no spatio-temporal relations, nothing that is good or bad, and no moral obligations—nor are we capable of conceiving of them. The noumenal world is one in which, insofar as we have a 'point of view' ('self', or 'transcendental ego') in it at all, is a world without problems or change.

The standpoint from which things are conceived *sub duratione*, or as in time (and place), is a standpoint that is both within and concerned with the phenomenal world. In the phenomenal world, according to Spinoza, there is "true knowledge of good and evil," although all such knowledge is 'abstract,' inadequate, and (as previously mentioned) literally inconceivable from the perspective of the noumenal world.

A central problem with Spinoza's account of the human mind, however, is that in the last part of the *Ethics*, he seems to suppose that it contains two quite disparate parts. In E V p29d he maintains that two things belong to the essence of the mind: (i) to conceive the essence of the human body *sub aeternitatis specie*; and (ii) to conceive the existence of the body *sub duratione*. Insofar as the mind conceives things in the first way, it is itself eternal—a noumenon, in fact—while insofar as it conceives them in the second way, it is a phenomenon in time. The eternal part of the mind is the intellect, Spinoza tells us, while the part that perishes is the imagination.[76] It is through the former that we act, and through the latter that we are passive.

[76] E V p40c.

But how a unity can be formed from two parts, one of which is atemporal and the other temporal, is not easy to see. It is no easier to understand how an atemporal intellect can be a partial cause of imaginative ideas that are in time.

Thus Spinoza's attempt to provide a unitary account of the human mind seems to encounter a problem quite like, if not the same as, Kant's difficulties with the distinction between noumena and phenomena.

Abbreviations

CM = *Cogitata Metaphysica* (*Metaphysical Thoughts*) This is the appendix to the *PPC*.
E = *Ethica* (*Ethics*).
Ep = Epistolae (Letters). These are numbered as in Spinoza (2000).
G = *Spinoza Opera*, ed. Carl Gebhardt (Heidelberg: Carl Winters Universitaetsbuchhandlung, 1925) ("G" is followed by volume (I–IV), page, and line numbers).
KV = *Korte Verhandeling van God, de Mensch en des zelfs Welstand* (*Short Treatise on God, Man, and His Well-Being*). "KV" is followed by part numbers (I or II), as well as chapter numbers.
PPC = Principles of Cartesian Philosophy, that is, *Renati Des Cartes Principiorum Philosophiae Pars I et II.* (*Part I and II of René Descartes' Principles of Philosophy*.) "PPC" is followed by part numbers (I, II, or III) and propositions, etc. as in E.
S = *Spinoza. Complete Works*, transl./ed. Samuel Shirley (Indianapolis/Cambridge: Hackett, 2002).
TdIE = *Tractatus de Intellectus Emendatione* (*Treatise on the Improvement of the Intellect*).
TP = *Tractatus Politicus* (*Political Treatise*).
TTP = *Tractatus Theologico-Politicus* (*Theologico-Political Treatise*).

Spinoza's Eternal Self

Olli Koistinen

Introduction

In this paper, my principal aim is to consider Spinoza's notions of the self and mind. It seems to be rather unanimously accepted, among Spinoza scholars, that Spinoza had a Humean bundle theory of the mind and the self. Any human self is just a bundle of God's ideas. Because of the problems faced by the bundle theory of the self, one should, as much as possible, avoid attributing such a view to any great philosopher of the past. In this paper, I will first consider the reasons why scholars are so fond of attributing the bundle view to Spinoza. I argue that it is true that Spinoza thought that any mind is a bundle of ideas but that bundling by itself is not sufficient for there to be a mind that belongs to somebody. After that, I will give a positive account of Spinoza's theory of the self. It will be claimed that, in a sense, any thinking thing is identical with God. Thus, the self of Charles Dickens is Spinoza's only substance God acting with a certain force characteristic to Charles Dickens. This idea will be further elucidated by Immanuel Kant's notion of intensive magnitudes. It will be claimed that once Spinoza's theory of the self is given this interpretation, there is room for individual immortality which should be distinguished from personal immortality. Also this view swims against the current of contemporary Spinoza research.

1 Bundle Theory

1.1 Subjects and Bundles: the Problem of Subjectless Bundles

The temptation to attribute a sort of bundle theory to Spinoza seems to stem from two principles. First, Spinoza thinks that individual bodies are composite entities consisting fundamentally of simplest bodies (*corporae simplicissimae*). When a set of such simplest bodies form a composite body they are tied to each other through a relation Spinoza calls ratio of motion and rest. It is this *ratio* which fundamentally is responsible for the individuality of a body. Parts of the same body may come and go without the destruction of that body if the same ratio of motion and rest is preserved. Second, Spinoza thinks that there is a perfect parallelism between mental and bodily items. Suppose now, that a body B exists

J. Miller (ed.), *Topics in Early Modern Philosophy of Mind*, Studies in the History of Philosophy of Mind 9, 151–169.

which at a certain moment is composed of two simplest bodies x and y. That x and y form a body means that they are tied to each other by a certain ratio R. Because of the parallelism there are mental counterparts, or ideas, of x and y, and also of the relation R, which will be abbreviated as I(x), I(y), and I(R) respectively. Thus, a mental individual I(B) whose individuality is determined by I(R) exists when B exists. Of course, there are several more or less sophisticated versions of Spinoza's bundle theory and this rough picture does not do full justice to them, but I believe the problems in this simple version are imported to the more sophisticated versions, too.[1]

It seems that the bundle theory is built on plausible premises, and I do not want to quarrel with them. Thus, I accept that there is a parallelism between compositional mental entities and bodies. But the problem is why these compositional mental entities that correspond to bodies should be minds of subjects; i.e., why they couldn't just be free-floating complexes of ideas. Moreover, why can't it be the case that these complex mental entities be complex ideas in one and the same mind without any multiplicity of subjects? (Of course, Spinoza allows that these complex ideas are all in God, but he grants the existence of several subjects.) So, the bundle theorist has to answer that the relation I(R) is of such nature that its instantiation is, *ipso facto*, the instantiation of a subject distinct from the instantiation of any other I(X). However, it seems difficult to hold such a position. To explicate, I do not want to deny that the mind of a human subject requires a certain kind of coherence of ideas or that her mind has to be united to a body in some way or other. But the big question that I believe has to be faced is what makes a complex of ideas, be it as coherent as it may, the mind of somebody. A particular human mind may contain extremely coherent subsets but it seems that it would be rather daring to call these subsets minds that belong to distinct subjects. One might be tempted to answer this by pointing out first that for Spinoza a body is individuated by its form, a ratio of motion and rest, and then, second, that for Spinoza there is a perfect parallelism between mind and body. Thus, corresponding to any individual body there is an individual mind. This is certainly true, but it would give no answer to the big question, i.e. what makes this mind the mind of someone.

1.2 Acting Bundles

It is extremely important for Spinoza that human beings act. Even though he stresses the finitude of human beings and their being constantly exposed to external causes by which they are affected, he also leaves room for certain events being completely caused by us and thus being actions. Spinoza defines acting as follows:

> I say that we act when something happens, in us or outside us, of which we are the adequate cause, i.e., (by D1), when something in us or outside us follows from our nature,

[1] For a version of bundle theory in Spinoza, see (Della Rocca 1996, pp. 41–43).

> which can be clearly and distinctly understood through it alone. On the other hand, I say that we are acted on when something happens in us, or something follows from our nature, of which we are only a partial cause.[2]

What this definition appears to say is that when human beings act, they are first or originating causes. They make something happen without being pushed by other things. That this is so follows directly from Spinoza's causal axiom, I a4, according to which an effect is conceived through its cause. Thus, if E's following from S's nature N is determined by an external cause X, then a complete understanding of E requires understanding of the nature of external cause X, too.

It seems to me to be very difficult to reconcile the bundle theory of subjects with the possibility of acting. The bundle that in this view is a human being is a finite bundle and is according to I p28 determined both to exist and to produce effects by other finite things. One might try to counter this objection by claiming that even though the whole human being is a finite mode of God, there are constituents of this bundle that are infinite and which partly help to constitute the nature of the human being. At first sight, this does not seem hopeless. Spinoza quite explicitly claims that we have adequate ideas and, moreover, he also holds that both adequate ideas and their effects are actions.[3] So, in spite of our finitude, we would succeed in acting when these adequate ideas are causally efficacious.

Without going into the details of this suggestion, I just want to point out that for Spinoza[4] all ideas belong to *natura naturata* and not to *natura naturans*, which means that they are caused by God and hence conceived through God; this also means that the adequate ideas of human beings should be conceived through God's nature, which suggests that human beings cannot act in having adequate ideas if human beings do not have an idea-free nature that somehow is identical with God's nature. Thus, our nature, which is causally responsible for the fact that we act, cannot be constituted by ideas of any sort, and our subjecthood cannot be reduced to ideas or to bundles of them.

In recent Spinoza scholarship, it has been argued that for Spinoza activity is always a matter of degree. Human beings may be more or less active and it is erroneous to see acting as an all-or-nothing matter. It is, of course, right to see Spinoza as holding that human beings are more or less active, but the degree of

---

[2] *Ethics* III d2. Translations are from Curley's *The Collected Works of Spinoza*, vol. 1. I have used the following method in referring to the *Ethics*: the first Roman number identifies the part of the *Ethics*, the letter after that specifies whether a proposition, p, definition, d, or axiom, a, is intended. The number after these letters tells which proposition, definition or axiom is referred to. If these numbers are followed by letters, then s means scholium and c corollary. For example, E I p16c1 refers to the first corollary of the sixteenth proposition of the first part of the *Ethics*.

[3] Adequate ideas have to be infinite because if they were not, they would not be constitutive to understanding, but would be like conclusions without premises. If an adequate idea were finite for Spinoza, it would have to exist in an infinite causal chain (E I p28). But because the knowledge of an effect involves knowledge of its cause, we would never reach adequate knowledge through finite ideas.

[4] See E I p31.

activity is dependent on amount and greatness of effects caused by the agent alone. Thus, human beings could not be active in any degree if they were not complete causes of some effects.

2 Substantial Self

The high standards Spinoza sets for acting seem to be inconsistent with his metaphysical system. It seems that what he says could be quite consistent with a Cartesian view of human beings where each human being has a substantial self that is capable of forming clear and distinct ideas and thus capable of acting. Spinoza's view of human beings as causally independent origins of action also resembles quite astonishingly Immanuel Kant's commitment to *actus originarius*.[5] For Kant such an action is an *a priori* action that is not determined by anything outside the agent. In his theoretical philosophy, these actions are mental actions of synthesis. Kant also thought that the knowledge we have of ourselves is through these original actions. There is no direct access to selves but our actions express ourselves. The basic feature of a self, as well as of any other substance, for Kant is the force through which the self acts.[6]

Because Spinoza was a substance monist, he believed in substance. The only substance, when considered under the attribute of thought, cannot be reduced to a bundle of ideas. It is something that is prior to these ideas and causes them, as has already been pointed out. This sole substance is also, according to Spinoza, the first cause of everything. Moreover, this only substance has an infinite force of existence and action which force is God's essence, through which everything is ultimately conceived. So how can there be, in such a system, where all actions are of God, a multitude of freely acting agents?

3 Self and God

3.1 Activity and Substance

The problems above seem to be generated by the assumption that there is some kind of distance between us and God, and thus it overlooks the very core of Spinoza's system that we are modifications of God. Thus, when Charles Dickens thinks this just means that God *qua* Charles Dickens thinks and God *qua* Charles Dickens is in Spinoza's ontology identical with the man Charles Dickens. So

[5] For a good discussion of self-determination in Spinoza, see Viljanen (2007, pp. 220–228).

[6] Of Kant's views, see Kant (1998, pp. B130–136).

when Charles Dickens engages in adequate thinking it is God who is thinking insofar as he is modified by a modification that results in the existence of Charles Dickens. Spinoza writes:

> [T]he human Mind is a part of the infinite intellect of God. Therefore, when we say that the human Mind perceives this or that, we are saying nothing but that God, not insofar as he is infinite, but insofar as he is explained through the nature of the human Mind, or insofar as he constitutes the essence of the human Mind, has this or that idea; and when we say that God has this or that idea, not only insofar as he constitutes the nature of the human Mind, but insofar as he also has the idea of another thing together with the human Mind, then we say that the human Mind perceives the thing only partially, or inadequately.[7]

This passage is central in the attempt to understand the relation between thinking things and God. After Bennett's (1984) treatment of monism about extended substance it has become understandable how bodies can be seen to be modifications of the attribute of extension. For example, a particular stone is fundamentally a region of space that is stone-like, as it were. Corresponding to any individual, there is a property associated with that individual and the individual is nothing but space's possessing that property. But how should this monism be understood on the mental side? The bundle theory as it is sketched above seems to offer no great help. No subset of God's ideas can be an active and self-determining thinking agent—help from substance is needed.

3.2 Essence of God and Essence of the Self

The essence of God, as Viljanen (2007) has convincingly argued, is infinite force and the attributes are expressions of this force. Moreover, force, or conatus or striving, also constitutes the essence of singular things and, in Spinoza's monism, the force of a singular thing is a portion of God's infinite force. The force by which a singular thing exists and acts follows, according to Spinoza, from God's essence, i.e., from his force. Spinoza writes:

> I am speaking, I say, of the very existence of singular things insofar as they are in God. For even if each one is determined by another singular thing to exist in a certain way, still the force [*vis*] by which each one perseveres in existing follows [*sequitur*] from the eternal necessity of God's nature.[8]

Once we grant that the essence of a singular is a portion of God's force, this passage just says that there is a certain kind of "whole-part" relation between the infinite force of God and the finite force of a singular thing. The infinite force of

[7] E II p11c.

[8] E II p45s.

God can be seen to be constituted by the finite forces of singular things.[9] What follows directly from God's essence are the essences of singular things, but the ways the singular things exist are determined by finite elements.

Let us now return to the II p11s passage quoted above, where Spinoza says that the human mind explains God's mind which he takes to be equivalent to saying that the human mind is constituted by God's mind. Spinoza is there speaking of the adequate part of the human mind, i.e., of the actions in the human mind. When one reads this passage along with the other passages where Spinoza claims that God's mind is explained by the human mind, and in conjunction with V p40—

> [O]ur Mind, insofar as it understands, is an eternal mode of thinking, which is determined by another eternal mode of thinking, and this again by another, and so on, to infinity; so that together, they all constitute God's eternal and infinite intellect[—]

the following picture emerges. There is an infinity of minds which determine each other but of which each one is eternal; together they constitute God's eternal and infinite intellect. Now, corresponding to each of these intellects there should be a specific force and these forces, in their turn, should constitute God's infinite force, i.e., God's essence. This follows because things act through their force and the extent of an intellect is determined by the force creating it. To understand what V p40 says is challenging but also very important for obtaining a picture of the relation between God's thought and our thought.

That God's intellect is explainable by the adequate parts of finite intellects seems to go against Spinoza's view of the explanatory priority of infinite with respect to anything that is finite. Moreover, it is hard to understand how God's infinite force could be seen to have a compositional structure; i.e., what Spinoza says may make one feel that somehow God's force should be composed of the finite forces not completely different from the way an aggregate is composed of its parts. In what follows, I will try to give an answer to these puzzles.

3.3 God's Intellect as a Compositional Entity

When Spinoza says in V p40 that God's infinite intellect is somehow constituted by the eternal intellects of particular things he faces a great difficulty. What V p40 suggests is that God's intellect is a compositional entity, but it is not at all easy to see how an intellect could be such an entity. From the fact that A and B are different intellects of two distinct subjects, it does not follow that there is a third intellect which consists of the ideas of both A and B. That you think something and I something else does not, of course, entail that there is something that thinks both what you think and what I think.

[9]Spinoza (2001) writes: "[T]he universal power of Nature is nothing but the power of all individual things taken together." (16.2)

It seems evident that Spinoza thought that somehow monism with regard to the thinking substance should make compositional intellects possible. Suppose I think A and you think B. In Spinoza's monism this entails that God thinks A and God thinks B. Thus, it should in principle be possible to see God's intellect as being constituted by the intellects of finite beings. However, as has already been argued, this rather blunt way of dealing with the problem seems to erase finite intellects from the picture. If my thinking A is *just* God's thinking A, then, because I am not identical with God, I do not think A. This seems to be *almost* an analytical truth. Of course, if by finite thought (i.e. the thinking done by finite beings) is meant only perception of ideas, then one might argue that all the ideas perceived by me are ideas caused and, in that way also thought, by God. But perception of ideas is not, for Spinoza, thinking. So what we are looking for is the possibility of a composite intellect without destroying the thinking subjects whose intellects are elements in the compositional intellect. It should be explicated how is it possible, as it were, that God thinks A by my thinking of A, so that both I and God are genuine thinkers of that thought.

Spinoza's leading thought seems to be that somehow all the (adequate) thinking there is going on is done by God and that all the thinking God does is done by the eternal intellects that determine each other, without it being the case, on the one hand, that God's intellect is reducible to the infinity of eternal intellects that determine each other, and, on the other hand, without it being the case that the infinity of thinking subjects is somehow illusory.

3.4 Thinking Force

I believe that a coherent picture of Spinoza's self-thinking can be reached once attention is paid to the nature of force, which, as has already been stated, is God's essence and from which the forces of singular things follow. As has also been claimed above, the forces of singular things are not really distinct from God's infinite force but are portions or constituents of that infinite force.

Giving a correct view of the nature of force is not quite easy. It seems promising to me to treat forces as intensive magnitudes, to use Kant's terminology, in order to distinguish them from extensive magnitudes, such as stones, which can be conceived to have parts that are external to each other. The quantity of force is given in degrees. The degree of the force F may be greater than the degree of the force G. For example, the intensity of light in this room may be greater than the intensity of light in the next room. By bringing these light sources into one room, one will reach a new intensity of light that, in principle, could be obtained by a light source that is more powerful than each of the previously mentioned two. Moreover, physical forces can be combined in such a way that several forces acting on a subject can be replaced by one force that is the resultant of the several distinct forces. For example, two forces which have intensities F and G and which

operate to the same direction on the body A, can be replaced by a force which has intensity F + G and which operates to the same direction.[10]

Let us now take seriously Spinoza's idea that thoughts are effects of force. Any intellect should thus be seen as an effect of force. When this is combined with the view that the intellect has its own order which proceeds from the first causes, it follows that there is no room for intellectual force composition that is analogous to the force composition in the physical world.[11] Any two intellects in this picture share the ideas of the first causes and they can differ only in their extent. This means that a greater intellect has the force to follow the order of the intellect longer than the smaller one. So, if we somehow combine two intellects the result cannot be anything else but one of the component intellects, provided it is not possible that two intellects share all their ideas. But Spinoza wants to see God's intellect as being composed of an infinity of eternal intellects which determine each other. What does that mean?

One suggestion could be the following. There exists an infinity of intellects, each belonging to a distinct thing down from the lowest to the greatest perfection. God's intellect, of course, is the most perfect intellect. Now, we could say that each of the imperfect intellects is a constituent of God's intellect because God's intellect contains any set of thoughts there can be. What is intended by this picture is that the principle of plenitude is applied to the intellects so that any possible intellect is actual. But even though Spinoza believed in the principle of plenitude he would not have accepted this because it is inconsistent with his monism. This kind of view requires that any member belonging to the infinity of the intellects has a separate force of thinking and thus there would be an infinity of thinking substances. Moreover, here it would be extremely metaphorical to say that the different intellects constitute the infinite intellect of God.

Another suggestion, and I believe it captures what Spinoza had in mind, is this. Once an infinite intellect, which is able to complete the order of the intellect, is given, all the other intellects are *ipso facto* given, as being embedded into the infinite intellect. By, or in, thinking with an infinite force of thought, the infinite substance also thinks with all possible degrees of force of which degrees each one is responsible for the generation of an intellect that is a constituent of the infinite intellect. Suppose the force needed for Spinoza's intellect to exist is SP. In thinking with the infinite force, God is thinking also so that it produces what SP produces. Thus, SP is in God's infinite force and we could say that Spinoza's self is God acting with the force SP, but we should bear in mind that God acting with the infinite force and God acting are not two really distinct actions of God. Spinoza's intellect cannot exist without God's infinite intellect and God's infinite intellect cannot exist without Spinoza's intellect. So, it is both true that Spinoza's

[10] Kant on intensive magnitudes, see, for example, Kant (1997, pp. 466–469).

[11] Of the order of the intellect in Spinoza, see E II p18s.

intellect follows from God's intellect and that Spinoza's intellect helps to constitute God's intellect.[12]

Force, as well as other intensive quantities, has an interesting ontology. It does not have such a part and whole structure that a force with a certain intensity could be seen to be composed of forces of lesser intensities—i.e., they are not aggregates. However, the ordinary notion of force seems to involve the idea that something can be added to a force and that something can be taken away from it. The intensity of light can become smaller and smaller and can finally be reduced to zero. So, let us suppose that the intensity of a certain light is L and it is diminished into L/2 and then again it gets an increase so that its intensity rises back to L. In this kind of situation we could say that in a certain sense L has parts which are jointly responsible for the illumination being what it is. God's infinite force is different from the intensity of light because there cannot be any changes in it and it is given all at once. But that does not exclude us from seeing it as having a similar compositional structure that we can attribute to the intensity of light or, say, pain and pleasure.

What, then, is a self for Spinoza? I would suggest that on the basis of what has been said above any self is, in a certain sense, identical with God. My intellect is generated through God acting with a force that individuates me; yours is generated through God acting with the force that individuates you. So *I am God insofar as God is considered as acting through the force that individuates me*, and *you are God insofar as God is conceived to act through the force that individuates you*.[13] But here it should be kept in mind that these acts are parts of the same act through which God forms the infinite idea of himself through his infinite force of thinking.[14] Moreover, because God for Spinoza is a simple substance the same is true of all selves, too.

3.5 Self and Body

In the material of the fifth part of the *Ethics*, where issues about the self are touched on, Spinoza is talking about "the mind's duration without relation to the

[12] The following alternative solution to the problem of intellect composition was suggested to me by the referee for this volume. All intellects are identical in the sense that they share all the ideas, but they differ in the what they conceive clearly. Thus, in a sense there is no problem of composition. However, this solution goes against the key passage from E V p 40 cited above. Moreover, I do not find any convincing textual evidence for that reading. And further, it seems to require that any mind has all possible adequate ideas which they need not conceive clearly. However, it seems to me that Spinoza cannot accept such unclear adequate ideas. Further, this would make all intellects infinite, and as I interpret Spinoza, only God's intellect is infinite.

[13] It should be emphasized that God's infinite force is given at once. So, God's infinite force is prior to the forces of finite things. Thus, God's infinite force is not *generated* through some kind of composition of finite forces.

[14] See E I p21.

body."[15] This could be seen as saying that he is speaking about the mind only through the attribute of thought without paying any attention to its connection to the body. However, much of what follows (especially V p21 and V p23) suggests that for Spinoza the mind is not completely dependent on the actually existing body but contains a core that is independent of it. This core Spinoza calls the intellect.[16] The ideas of the intellect are our actions and so it is this active part of us that is independent of the body. The ideas that are dependent on the body are ideas of imagination on whose existence memory depends.

If I am correct about Spinoza's view of the self as identical with God insofar as he is conceived to act through a determinate force, thus being responsible for the being of an intellect whose existence is not dependent on an actually existing body, it follows that the self is not ontologically dependent on any actually existing body. This seems to bring Spinoza's theory of the self close to the one Descartes was holding, which, of course, is something that must cause some embarrassment among Spinoza scholars.

The prevalent view among contemporary Spinoza scholars about the eternal part of the human mind, i.e., about the intellect, is that it is just a part of the human mind which, considered in itself, is not necessarily attached to any particular self. When Spinoza writes at V p23 that

> [t]he human Mind cannot be absolutely destroyed with the Body, but something of it remains which is eternal[,]

this should be interpreted as meaning only that the adequate ideas somebody had in this life just continue their existence without being tied to the self whose ideas they were in this life. However, in the interpretation put forward in this paper, Spinoza's immortality doctrine should be seen as involving a commitment to individual immortality. In the following section, I will try to justify this reading.[17]

4 Individual Immortality and Bodiless Selves

Once a wedge is drawn between durational mind and body on the one hand and the self on the other hand there is room for understanding Spinoza's self as a non-bodily thing. But, on the contrary, it is essential to the self that it thinks. Anything that exists has to operate in some way and for Spinoza the essential activity of selves is, at least, thinking. If there are selves, these selves think.

The possibility of there being thinking selves that can be conceived to exist without their actual bodies creates a difficult problem. Spinoza thought that there

[15] E V p20s.

[16] E V p40c.

[17] For a good discussion of immortality in Spinoza and of different positions scholars have taken on that issue see Nadler (2001, Ch. 5).

is a perfect parallelism or some kind of identity between the ideas and their objects. So how can a thinking self exist without its body? It should be first noted that this is a general difficulty for Spinoza and not something that is tied to the present interpretation. Something, says Spinoza, remains of our mind after the bodily destruction. But how can that be if the objects of the ideas have ceased to exist?

The answer to the problem just presented can be approached by paying attention to two senses of existence in Spinoza: durational and eternal existence. Those ideas that have as their objects existing bodies do have durational existence. Thus, the durational existence of an idea is explicated in terms of the durational existence of its object. Such a durational object is constantly affected by other things,[18] and an idea which has a durational body as its object has actual being. However, not all ideas have durational objects and thus not all ideas have actual, or temporal, being. Those ideas which do not have actual being have as their objects formal essences; and for humans those formal essences are ideas of pure extension. So, there is in God an eternal idea of the human body under a form of eternity and this idea is the idea of the formal essence of the body. What remains of my mind after my bodily destruction is, then, the idea of the formal essence of my body.[19]

It seems that the prevalent tone in Spinoza scholarship is just to say that this is what my immortality is—the eternal existence of the idea of the formal essence of my body that is thought eternally by God. This seems a rather mild form of immortality which *prima facie* does not differ much from some photographs of me surviving my death. Was Spinoza's conception of immortality just that?

In fact, it seems to me that there is rather strong evidence for the opposite view. First, as has already been stated, Spinoza claims that the force by which a singular thing exists follows from the eternal essence of God. This suggests that the force, which is my essence, is eternally posited.[20] But my essence being eternally posited should mean nothing but that I am eternal. Second, at V p23s Spinoza writes:

> And though it is impossible that we should recollect that we existed before the Body—since there cannot be any traces of this in the body, and eternity can neither be defined by time nor have any relation to time—still, we feel and know by experience that we are eternal. For the Mind feels those things that it conceives in understanding no less than those it has in the memory. For the eyes of the mind, by which it sees and observes things, are the demonstrations themselves.

Of particular importance in this passage is the claim that we cannot recollect that we existed before the body, but that we feel and know by experience that we are eternal. It seems that in this passage it is rather directly claimed that we as subjects are eternal. What is further argued is that this eternity of ours cannot be justified

[18] E II p11, 2a4, 2a1.

[19] See Garrett (2008). I will consider formal essences more fully in Section 6.

[20] Spinoza identifies force (or conatus) with the actual essence of finite things at E III p7.

by memory but is demonstrated, and this demonstration gives us the feeling of its certainty—maybe in the way that demonstrating the Pythagorean theorem makes me feel that it is true. I believe that those supporting the anti-individual reading should say that by "we" Spinoza just means "part of us." Maybe, but it is hard to believe that at this most important place Spinoza would have been so careless in formulating his views.

Spinoza argues at V p38 that the more the mind understands things by the second and third kinds of knowledge, the less the mind fears death. The proposition is demonstrated in such a way that it is consistent with the individual and the anti-individual readings of Spinoza on immortality. The demonstration says that the more the mind understands, the less it is bothered by emotions that are harmful and fear is a harmful emotion. So, if somebody does not think of death, she doesn't fear it.

However, one wonders why a rational man would not have an attitude that is similar to the fear of death, if the anti-individual reading is accepted. Spinoza defines fear as follows:

> Fear, on the other hand, is an inconstant Sadness, which has also arisen from the image of a doubtful thing.[21]

In the anti-individual reading of immortality, our attitude towards our necessary individual death would not be fear because according to that reading there is no doubt about death. However, there is a still worse emotion than fear which arises when one is certain of the bad outcome. This is called despair:

> Next, if the doubt involved in these affects is removed, Hope becomes Confidence, and Fear, Despair—viz. a Joy or Sadness which has arisen from the image of a thing we feared or hoped for.[22]

Naturally, those supporting the anti-individual reading do not want to claim that a wise man does not fear death but is in despair about it, even though that seems to be what they should hold if individual immortality is denied. It seems to me that according to the anti-individual reading human beings are finite modes of God and death will destroy them so that no self remains. But it seems to me that even if something of the mind remained after bodily death, this should have no consolatory value for he who dies—the adequate ideas somebody has been thinking have been there for all eternity and for somebody who dies, the fact that he was able for a while to think a couple of adequate ideas should give no relief. Understanding in itself is a good thing and it is linked to intellectual pleasure. I feel pleasure while thinking adequately and this pleasure is intellectual pleasure. But suppose now, that I am certain that I die—I cannot know, Spinoza thinks, adequately when but I am certain it will happen. So why wouldn't I be in despair about it? My thinking will be terminated; there is no self that remains and nothing would be good to me. In fact, it seems that those people who have enjoyed life here very much fear death

[21] E III p18s2.

[22] E III p18s2.

more than those who have suffered several tragedies and who are mostly in pain. So why is death not fearful to active thinkers?

One might try to answer this by simply claiming that if in death the subject ceases to exist, it does not literally lose anything, or, to use Spinoza's terminology, the subject does not pass from a greater to a lesser perfection, and thus does not feel pain. In the same way, as there is nothing negative to the subject in the world before her so there is nothing negative, and therefore nothing to be feared, in the world after her. But this is not the way Spinoza argues. After having realized that death is not a loss to the subject, anyone could rid herself of the fear death irrespective of how much adequate thinking she has been exercising. But it seems that part of what Spinoza wants to say is that the more somebody thinks adequately, the less she will lose after the destruction of her body, and thus the loss is smaller than for somebody whose thinking is based on imagination.

It may well be that the passages where the fear of death is considered can also be treated equally well in the anti-individual reading, but I believe that the issue about the rationality of morality is more difficult to it. This may be surprising but let me explain what I mean. It is well known that Spinoza answered to the rationality of morality in a way that is hedonistic or at least compatible with a sort of hedonism. Only one thing seems to intrinsically good for Spinoza, namely understanding, and those things that help us to understand are derivatively good. Understanding or adequate thinking is something that necessarily involves feeling pleasure and it seems that the main motivation for its being good is just this essential connection with pleasure or joy. Thus, acting morally, i.e., doing those things that contribute to an increase in understanding, is worthwhile regardless of any possible reward in the hereafter.

Spinoza formulates the thought that moral action or acting in accordance with virtue need not to be justified through possible rewards after bodily destruction in an interesting way:

> Even if we did not know that our Mind is eternal, we would still regard as of the first importance Morality, Religion, and absolutely all the things we have shown (in Part IV) to be related to Tenacity and Nobility.
>
> Dem.: The first and only foundation of virtue, or of the method of living rightly (by IV P22C and P24) is the seeking of our own advantage. But to determine what reason prescribes as useful, we took no account of the eternity of the Mind, which we only came to know in the Fifth Part. Therefore, though we did not know then that the Mind is eternal, we still regarded as of the first importance the things we showed to be related to Tenacity and Nobility. And so, even if we also did not know this now, we would still regard as of the first importance the same rules of reason, q.e.d.[23]

The first point here seems to be that moral action need not be motivated by the thought that it makes us more eternal than what we would be if our lives were based on imagination. However, the second point which is suggested by the 'even if' (*quamvis*) phrase is that the connection between moral life and mind-eternity

[23] E V p41.

also motivates those who are aware of this to moral action. I do not see how mind-eternity, when it is interpreted in an anti-individual way, could provide such an extra incentive for moral action. What remains after my bodily destruction has, according to the anti-individual interpretation, no relation to me. (I don't want to say that all moral action should be based on egoistical considerations, but it is hard to find anything else from Spinoza.)

I will now turn to consider the relation between acting and self-consciousness. According to Spinoza our eternity is connected to acting:

> For the eternal part of the Mind (by P23 and P29) is the intellect, through which alone we are said to act (by III P3). But what we have shown to perish is the imagination (by P21), through which alone we are said to be acted on (by III P3 and the gen. Def. Aff.). So (by P40), the intellect, however extensive it is, is more perfect than the imagination, q.e.d.[24]

What seems to be the crucial question is whether this kind of eternal acting is connected somehow to an eternal self. I believe Spinoza thinks that it is because for him all adequate cognition involves adequate knowledge about one's having that cognition. In II p43 Spinoza writes:

> He who has a true idea at the same time knows that he has a true idea, and cannot doubt the truth of the thing.

As the demonstration makes clear Spinoza here identifies a true idea with an adequate idea, and so this proposition says that if somebody has an adequate idea then she knows that she has an adequate idea etc. So if S has an adequate idea, then S knows that she has an adequate idea. Thus, the subject herself is part of the content of the second-order idea. The demonstration, moreover, makes it clear that the second order idea is in itself an adequate idea. But suppose now that the subject which is involved in the second order idea is a durational subject or a finite thing which comes into being and perishes. Of such a thing we cannot have adequate knowledge; we know our durational mind and our durational body only through the affections of the body, and it seems to be evident that if an idea of such a thing is involved by some other idea, the complex idea itself is an inadequate idea. So it seems that we are forced to conclude that through adequate thinking we have (adequate) knowledge of a non-durational, i.e., of an eternal, self. Spinoza himself connects II p43 to self-knowledge at V p27d:

> [H]e who knows things by this [third] kind of knowledge passes to the greatest human perfection, and consequently (by Def. Aff. II), is affected with the greatest Joy, accompanied (by II P43) by the idea of himself and his virtue.

The importance of this passage is not limited solely to the fact that Spinoza really does connect II p43 to self-knowledge but that he also sees that virtue is connected to the idea of the self. Virtue for Spinoza is identical with force or power[25] and so self-knowledge is allied—as one might assume from the interpretation at hand,

[24] E V p40c.

[25] E IV d8.

where power or force is seen as the essence of the eternal self—with knowledge or idea of force.

There is still one line of thought that gives confirmation to the individual reading of immortality in Spinoza. In the scholium to the last proposition of the *Ethics*, Spinoza writes:

> For not only is the ignorant man troubled in many ways by external causes, and unable ever to possess true peace of mind, but he also lives as if he knew neither himself, nor God, nor things; and as soon as he ceases to be acted on, he ceases to be. On the other hand, the wise man, insofar as he is considered as such, is hardly troubled in spirit, but being, by a certain eternal necessity, conscious of himself, and of God, and of things, he never ceases to be, but always possesses true peace of mind.[26]

In this passage, Spinoza clearly ties the adequate third kind of knowledge to self-knowledge, or consciousness of oneself to eternal existence. But this passage may also be seen as generating a problem to my interpretation. It is said that the ignorant man ceases to be when he is no more acted on, whereas my interpretation would predict the eternal existence of all thinking selves. First, it should be said that a completely inactive human being cannot be anything but a negative ideal for Spinoza in the same sense as a completely self-determining human being is an unrealizable positive ideal. All human beings act by virtue of having certain adequate ideas. And thus, all human beings are immortal with respect to that kind of acting. However, self-knowledge or self-consciousness of passive human beings is continuously obscured by the passive ideas of imagination. But second, it seems that for Spinoza existence, or at least, reality allows for degrees. The more active somebody is, the more force she has, and thus, in a certain sense, more existence. It is, I believe, not misleading to say that there is a resemblance between the ways Descartes and Spinoza connect thinking to the existence of selves. "I think, I am" would for Spinoza become "The more I think, the more I am."

5 The Relation of the Self to Its Finite Mind and Finite Body

In the beginning of the second part of the *Ethics*, Spinoza considers the origin of the human mind. Here he is not concerned with the eternal aspect of us but is considering us as passionate, embodied and thinking beings—i.e., as men. The main aim in this difficult but important part of the *Ethics*, i.e., in propositions II p1–p13, is to give an explanation of the union of human mind and body. He writes as follows:

> From these [propositions] we understand not only that the human Mind is united to the Body, but also what should be understood by the union of Mind and Body.[27]

[26] E V p42.

[27] E II p13s.

To understand the union between the temporal mind and the temporal body one should somehow understand the cause, or generation, of the human mind, and for Spinoza this kind of union is generated by what I call object-taking.

"The first thing that constitutes the actual being of a human Mind is nothing but the idea of a singular thing which actually exists," writes Spinoza.[28] Before this Spinoza believes himself to have shown that thought is an attribute of God, which entails that there is only one substance and, moreover, that this substance thinks of everything that is possible. So when a particular human mind begins its actual existence it has as its object something that actually exists, and such an object cannot be anything else but the thing we call our body. That it cannot be anything else but our body is demonstrated very interestingly. Spinoza claims that God's knowledge of what happens in a singular thing, or in a singular object of any idea, is possible only through God's having an idea of that very same object. So, if there is in God knowledge about something happening in a body, that body has to be the object of the idea. God, then, cannot have any knowledge of what happens in other bodies indirectly. But how is this related to me, my mind and my body? No self has been included and the demonstration of II p13—viz., that "[t]he object of the idea constituting the human Mind is the Body, or a certain mode of Extension which actually exists, and nothing else"—takes the reader by surprise:

> For if the object of the human Mind were not the Body, the ideas of the affections of the Body would not be in God (by P9C) insofar as he constituted our Mind, but insofar as he constituted the mind of another thing, i.e., (by P11C), the ideas of the affections of the Body would not be in our Mind; but (by A4) we have ideas of the affections of the body. Therefore, the object of the idea that constitutes the human Mind is the Body, and it (by P11) actually exists.

The surprising element is that the fact that we feel a body as being affected is given such an important role in the demonstration. How is that fact supposed to function as a premise when the other premises refer to God's knowledge about what is happening in singular objects? So both I and God have ideas of affections of a certain body. But because all ideas there are, are God's ideas, it should follow that my having an idea of that body being affected is identical to God's having that idea. However, this can be the case only if I am in a certain sense identical with God—i.e., that God has that idea by virtue of my having that idea.[29] But given what has been said above, this makes perfect sense. I am God acting with a certain force of thinking; by coming to think of an actually existing body, through the common order of nature, my temporal mind and body begin to exist.

---

[28] E II p11.

[29] It would also be true to say that I have some idea by virtue of God's having that idea. There is no contradiction in holding both that God's thinking is a result of finite things' thinking and that finite things' thinking is a result of God's thinking. In the former alternative, explanation goes from "parts" to the whole whereas in the latter, from whole to "parts". In order to leave room both for genuine finite thinkers and for God as a thinking thing, this, of course, is how it should be.

6 On the Objects of the Minds of Eternal Selves

In my interpretation all particular human intellects are eternal and the subject of a particular intellect is God *qua* a force F where F is a portion of God's infinite force. This infinite force is what is needed for God to exist and in this way it is prior to the finite forces of individual subjects. Moreover, these eternal determinate subjects are, in my interpretation, thinking subjects which have an eternal intellect that also is part of the mind which has a limited duration. For Spinoza minds are ideas—possibly complex ones—which, by conceptual necessity, have objects. So what kind of objects do the minds of these eternal selves have?

To answer the question just posed, it is advisable, as always with Spinoza, to proceed from top to bottom. With the help of (i) his necessitarianism, which entails that everything that is possible is "actual," and (ii) that an infinite thinking being can think everything that is possible, Spinoza proves that (iii) there has to be in God an eternal idea of himself and of everything that follows from his essence; i.e., an idea of everything there is. This idea involves, as its constituents, the ideas of all bodies. Thus, there is in God necessarily the idea of my body. However, this idea does not have the durational body as its object but something else that Spinoza identifies as the formal essence of my body. These formal essences are eternal and as the demonstration of the eternity of the mind makes evident, the formal essence of the body gives content to the eternal part of a particular human mind.[30]

One wonders why the idea of the formal essences of my body should be included in the human mind. If the formal essence of the body is something like the general characterization of the human body, then, even if such a characterization exists in God's mind, why should it also exist in that particular mind which has as its temporal object a body with that formal essence? In order to answer this question, it is necessary to take a look at Spinoza's theory of essence, because, as Don Garrett (2008) has made clear, when Spinoza speaks about essences, he quite often means formal essences. Spinoza's definition of what pertains to the essence of a thing runs as follows:

> I say that to the essence of any thing belongs [*pertinere*] that which, being given, the thing is [NS: also] necessarily posited and which, being taken away, the thing is necessarily [NS: also] taken away; or that without which the thing can neither be nor be conceived, and which can neither be nor be conceived without the thing.[31]

This definition has been much discussed but here I would like to focus only on one element in it, namely on Spinoza's view that without the essence a thing can not be conceived. On one reading this would mean that if E pertains to the essence of x, then the proposition " x is not E" is inconceivable. This way of reading "without the essence a thing can not be conceived" does not rule out the

[30] Of the role of formal essences in Spinoza's treatment of immortality, see Garrett (2008).

[31] E II d2.

possibility of forming an idea of x without E and it would be compatible with the view that the idea of the formal essence of my body need not be in my mind. However, in another reading something cannot be conceived without the essence, if it is impossible to form any idea of a thing without the idea of its essence. And I believe that if we adopt this alternative, it becomes understandable why the idea of the formal essence of my body has to be an eternal element of my mind.

For Spinoza, as we have already seen, the object of the idea that constitutes the actual being of the human mind is the idea of an actually existing body. This idea is formed by God and, in my interpretation, by the self whose temporal being is constituted by this kind of object taking. But how does God, or I, succeed in thinking of my body if my body has an essence without which it cannot be conceived at all? It seems that a serious alternative is that the idea of this essence exists prior to the object-taking in God's mind insofar as he is explained through the nature of my mind. Thus, the idea of my body, i.e., my mind, involves the idea of the formal essence of my body which, by II p3, has to be eternally in God's mind.

In general, Spinoza is friendly to the view that there are *a priori* requirements for thinking about objects. We cannot have ideas of bodies without having the adequate idea of God's extension.[32] But as an adequate idea it cannot be sensory based and, for that reason, it is an *a priori* condition of our having thoughts about bodies at all. In a similar way, the idea of the formal essence of my body should be seen as an *a priori* condition for my mind's being necessarily a mind that can have only a human body as its object.

It is still a question worth pondering whether the formal essence completely individuates my body or whether there is some generality in it. The following passage from V p39 suggests that the formal essence is not tied to a particular body:

> In this life, then, we strive especially that the infant's Body may change (as much as its nature allows and assists) into another, capable of a great many things and related to a Mind very much conscious of itself, of God, and of things. We strive, that is, that whatever is related to its memory or imagination is of hardly any moment in relation to the intellect (as I have already said in P38S).

This passage allows that the body of infancy may be changed into another one. However, this change has limits to it. The body of infancy cannot be changed into a flower, for example. One way of explaining this is to claim that the human mind cannot have as its direct object anything else but the human body because the human mind can take as its objects only those objects whose formal essences are already in the human mind. And as Spinoza holds, the formal essences involved in human minds are formal essences of bodies. So, I would suggest that formal essences cannot individuate a body.

[32] See also Garrett (2008).

Conclusion

In this paper, I have argued that Spinoza did not have a bundle theory of the self. His basic idea was that a self is God acting with the force that individuates that self. It was suggested that once Spinoza is interpreted in this way, the problem of composite intellects posed by V p40s can be solved. Moreover, it has been claimed that in Spinoza's system there is room for individual immortality which should strictly be kept separate from personal immortality which includes memory and imagination. However, in my interpretation there is a very tight connection between ourselves and human bodies. It is not just that it happens through the common order of nature that our temporal realization or maybe realizations are tied to human body. Any eternal self has an eternal mind whose essential component is the idea of the formal essence of the body. This formal essence of the body is an *a priori* condition for our being able to form an idea of an actually existing body. Such an object-taking, as I have called it, is necessary for our existence as human beings which are constituted by the union of mind and body. But because there is no room in our eternal minds for ideas of formal essences of objects that belong to other attributes than to that of extension, we can have durational existence only as human beings.

Can Matter Think? The Mind–Body Problem in the Clarke–Collins Correspondence

Marleen Rozemond

Introduction

Descartes' mind–body dualism is frequently thought to lie at the origins of our concerns about the relationship of the mental to the physical. His approach to the issue is very different from that of most current philosophers and he is frequently a target for their criticism, as contemporary philosophers tend not to be dualists but are generally more inclined towards some form or other of materialism. There is, however, another significant difference between Descartes and current discussions: unlike contemporary philosophers, Descartes focused on arguing for *substance* dualism. The question he addressed at length was the question whether thinking and material qualities could belong to the same substance. He thought it pretty obvious that thinking is not identical with motion or other material qualities[1] and he never addressed other ways in which thinking or consciousness might fail to be a metaphysically fundamental category distinct from material qualities. Contemporary philosophers, however, focus on the relationship between mental and physical *states*; the classical notion of substance has disappeared from the scene.

The question whether we can establish the immateriality—and immortality—of the human soul continued to be very important in the early modern period after Descartes, and other early moderns, such as Locke and Leibniz, also tended to focus on substance dualism. A striking exception is the correspondence between Samuel Clarke, best known for his correspondence with Leibniz, and Anthony Collins.[2] Collins was a freethinker, materialist, and deist well-known at the time in England, who was close to Locke during Locke's later years. Their exchange, which took place in 1706–1708, devotes extensive attention to the question whether mental states are a metaphysically fundamental category distinct from physical states, or might arise from or be identical to physical states. The correspondence

[1] Descartes did not think he needed to *argue* that thinking is not motion; if one withdraws from the senses and uses one's intellect properly it should be obvious. See the Sixth Replies to the *Meditations* (AT VII.425,441, CSM II.287,297). Reference to Descartes' writings are as follows: AT: Charles Adam and Paul Tannery eds., *Oeurves de Descartes*, 11 vols. (Paris: CNSR and Vrin: 1964–1976); CSM: *The Philosophical Writings of Descartes*, 3 vols. (Cambridge: Cambridge University Press, 1985–1991).

[2] The correspondence can be found in Samuel Clarke, *The Works* (henceforth W), vol. III (London: 1738; reprint, New York: Garland Publishing, 1978). Italics are in the original text.

J. Miller (ed.), *Topics in Early Modern Philosophy of Mind*, Studies in the History of Philosophy of Mind 9, 171–192.

was part of the thinking matter debate ignited by Locke. Unlike Descartes, Locke thought that substance dualism cannot be established, because he thought that we cannot rule out the possibility that God superadds thinking to matter:

> We have *Ideas* of *Matter* and *Thinking*, but possibly shall never be able to know, whether any mere material Being thinks, or no; it being impossible for us, by the contemplation of our own *Ideas*, without revelation, to discover whether Omnipotency has not given to some System of Matter fitly disposed, a power to perceive and think, or else joined and fixed to Matter so disposed, a thinking immaterial Substance ….[3]

Locke's presence can be felt keenly in the correspondence: indeed, both interlocutors invoke various claims of his to support their own.

The heated debate about thinking matter occupied a number of thinkers on both sides of the English Channel over the course of the next century. It has received relatively little attention from historians of philosophy, in spite of its importance at the time, and in spite of the prominence of the mind–body problem in contemporary Anglo-American philosophy.[4] A reason may be that much of the discussion was carried out by philosophers who tend to receive less attention than the canonical figures in this period—such as Berkeley, Leibniz or Hume.

The Clarke–Collins correspondence was prominent in this debate. It started off with a public letter by Clarke in response to a book by Henry Dodwell who argued that the soul is not naturally, but only supernaturally immortal: God makes it continue to exist after death. Clarke objected and argued that the soul is immaterial and naturally immortal. It set off a public correspondence with Collins, who took Dodwell's side. The collected letters went through six editions, and was discussed in at least Britain and Amsterdam throughout much of the 18th century.[5] Leibniz received the correspondence, and commented that he thought Clarke made some good points, while disagreeing with others and thus the Leibniz–Clarke correspondence was ignited.

Clarke makes very clear that his ultimate concern with the issue of thinking matter lies in traditional religious questions about the immateriality and immortality of the soul. In his initial letter to Dodwell, Clarke claimed that an appropriate view of the afterlife is that the human soul is naturally immortal, that is, it cannot go out of existence as a result of natural processes, and this means that it cannot be

---

[3] Locke, *An Essay Concerning Human Understanding* (henceforth *Essay*) IV.iii.6, pp. 540–541. Locke did think that we can establish that God is an immaterial thinking substance. See *Essay* IV.x. When the quotes contain italics, they are in the original texts. For emphasis I have added underlining.

[4] But see John W. Yolton (1983), *Thinking Matter: Materialism in Eighteenth-Century Britain*, (Minneapolis: University of Minnesota Press, 1983), and *Locke and French Materialism* (Oxford: Oxford University Press, 1991).

[5] Robin Attfield (1977), "Clarke, Collins and Compounds," *Journal of the History of Philosophy* 15: 45–54, esp. p. 47.

material.[6] Later in the correspondence Clarke charges that materialism is a problem for religion because it threatens free will, opens the gates to believing that all rational beings are material, including God, and raises serious problems for the afterlife and "the Justice of future Rewards and Punishments."[7]

Clarke begins with an argument for the immateriality of the human soul that is a version of what Kant in the Second Paralogism called the "Achilles of dialectical inferences in the pure doctrine of the soul."[8] This is an old argument; it has roots in Plato and it can be found as early as Plotinus. It enjoyed considerable popularity in the early modern period.[9] Taking my cue from Kant, I will speak of the Achilles Argument. The argument is run variously in terms of mental activity generally speaking or particular types of mental processes, and it contends that a mental subject must be simple. In Clarke's version, the argument contends that consciousness or thinking requires that it belong to an "individual being." Otherwise consciousness would be the sum of consciousnesses of the parts, and this is not possible. He argued that matter cannot constitute such a being. Collins agreed that consciousness could not belong to a material subject in virtue of the parts of such a subject being conscious, but he argued that thinking could belong to a material subject in other ways. Perhaps his most interesting response is that thinking could result from, or as we might now say, emerge from, material qualities that characterize the parts of the system of matter. In this paper I will focus on their discussion of emergentism, which takes up the bulk of their discussion of the possibility of thinking matter.

I will first briefly discuss Clarke's Achilles Argument, then I will turn to Collins' proposal of emergentism. I will then examine Clarke's rejection of emergentism, which centers on what I will call The Homogeneity Principle (HP), according to which a quality of a composite whole must be the "sum and result" of qualities of the parts, and those qualities must be of the same kind as the quality of the whole in question.[10] Clarke argues that this principle applies to what he calls "really inherent qualities," and consciousness is one of those. Collins offers counterexamples to the HP but Clarke argues that they are not examples of really inherent qualities. In the end, Collins accepts a suitably specific version of HP, but then proposes that consciousness could be identical with a mode of a material

---

[6] While for Clarke the soul is naturally immortal, he thinks its existence always depends on God, who can annihilate it at any time—a view commonly held in the period about all creatures (W III.722).

[7] W III.851.

[8] *Critique of Pure Reason* A 351.

[9] For an extensive history of the argument, see Ben Lazare Mijuskovic, *The Achilles of Rationalist Arguments. The Simplicity, Unity, and Identity of Thought and Soul from the Cambridge Platonists to Kant: A Study in the History of an Argument* (The Hague: Martinus Nijhof, 1974).

[10] I owe the term to Ezio Vailiati, who speaks of the principle of homogeneity. Vailati offers some discussion of Clarke and Collins' exchange about thinking matter in his (1993) "Clarke's Extended Soul," *Journal of the History of Philosophy* 31: 387–403; and *Leibniz and Clarke: A Study of Their Correspondence* (Oxford: Oxford University Press, 1997), pp. 53–77.

quality. At that point the discussion shifts from emergentism to the possibility of identity between consciousness and material qualities. I conclude with a discussion of a disagreement between Clarke and Collins that runs through their exchange about the scope of our knowledge: Clarke is confident that this knowledge is sufficient to rule out various forms of materialism, Collins disagrees.

This correspondence is rather obscure, but it is of special philosophical interest, given its detailed investigation, unusual for the period, into the question whether thinking or consciousness can be identical with or emerge from material qualities. The correspondence offers the possibility of a deeper understanding of how at least some philosophers in the period thought about this aspect of the mind–body problem. The following conclusions are suggested by the exchange between Clarke and Collins. First, Collins proposes emergentism as a way of avoiding the problem of consciousness of a complex subject being the sum of the consciousnesses of the parts. This suggests that the success of the Achilles Argument requires that thinking or consciousness cannot emerge from material qualities. Second, the main obstacle to emergentism turns out to be a type of constraint on causation that was widely accepted during the period. Third, an important disagreement between Clarke and Collins concerns the scope of our knowledge, a disagreement that separated many philosophers in the period, and that turns out to bear significantly on their disagreement about the mind–body problem. Although I will not be able to explore this angle, the discussion of emergentism has clear resonances with current discussions of the mind–body problem.

Before we begin, a few remarks: Clarke and Collins talk about qualities or powers interchangeably, as did Locke. As is typical in this period, they are talking about particular instances of qualities, what nowadays are called tropes: the particular instances of consciousness that belong to a mind, the particular instances of shape, size and motion that can be found in a particular body. Secondly, the discussion takes place in the context of an early modern mechanistic conception of bodies. So the kinds of material qualities that are assumed are shape, size, motion. The status of gravity is explicitly part of what is at stake, and Clarke and Collins disagree about it.

1 Substance Dualism and the Achilles Argument

Clarke states the Achilles Argument as follows:

> For *Matter* being a divisible Substance, consisting always of separable, nay of actually separate and distinct parts, ‘tis plain, that unless it were essentially Conscious, in which case every particle of Matter must consist of innumerable separate and distinct Consciousnesses, no system of it in any possible Composition or Division, can be any individual Conscious Being; For, suppose three or three hundred Particles of Matter, at a Mile or any given distance one from another; is it possible that all those separate parts should in that State be one individual Conscious Being? Suppose then all these particles

> brought together into one System, so as to touch one another; will they thereby, or by any Motion or Composition whatsoever, become any whit less truly distinct Beings, than they were when at the greatest distance? How then can their being disposed in any possible System, make them one individual conscious Being? If you will suppose God by his infinite Power superadding Consciousness to the united Particles, yet still those Particles being really and necessarily as distinct Beings as ever, cannot be themselves the Subject in which that individual Consciousness inheres, but the Consciousness can only be superadded by the addition of Something, which in all the Particles must still it self be but one individual Being.[11]

Three points about this argument.

(i) In light of their later disagreements, it is worth noting that Clarke displays clear affinities with Leibniz, who expressed his approval of this argument.[12] Leibniz too held that the subject of perception cannot be material because perception requires a simple subject.[13] Furthermore, Clarke's contention that bringing particles of matter together won't help generate a genuine individual echoes a similar argument Leibniz offered. He too held that matter is essentially lacking in unity. Consider two diamonds, he wrote to Arnauld; when they are separated in space they are not one being. If we bring them close together, even if they are set in the same ring, they still do not constitute a single substance.[14]

(ii) This last point about the nature of matter is very important to the argument: like Leibniz, Clarke thought that matter cannot constitute a genuine individual, for matter is always an aggregate consisting of actually distinct parts. In Clarke's terms, a piece of matter could never be an "individual" subject, and it could never have the type of unity requisite for a subject of consciousness. This is a view Collins questions. The issue of the nature of matter in the correspondence, however, is a subject for another time.[15]

(iii) It is tempting to see Clarke's argument as a unity of consciousness argument of the kind Kant discusses in the Second Paralogism. But it is worth

---

[11] Clarke, W III, p. 730.

[12] André Robinet, *Correspondence Leibniz-Clarke; présentée d'après les manuscrits originaux des bibliothèques de Hanovre et de Londres* (Paris: Presses Universitaires de France, 1957), p. 21.

[13] *Monadology* XVII. For discussion of these issues in Leibniz see Margaret Wilson, "Leibniz and Materialism" in *Ideas and Mechanism: Essays on Early Modern Philosophy* (Princeton: Princeton University Press, 1999); and Marc Bobro and Paul Lodge (1998), "Stepping Back Inside Leibniz's Mill," *The Monist* 81: 554–573.

[14] G II.76, AG 79. For references to Leibniz' work in the original languages see *Die Philosophischen Schriften von Gottfried Wilhelm Leibniz*, ed. C.I. Gerhardt, 7 vols. (Berlin, Wiedmann, 1875–1890, repr. Hildesheim, Georg Olms, 1978) (G). Translations can be found in G.W. Leibniz, *Philosophical Essays*, eds. Roger Ariew and Daniel Garber (Indianapolis, IN: Hackett, 1989) (AG).

[15] I discuss this issue at length in "The Achilles Argument and the Nature of Matter in the Clarke–Collins Correspondence," in *The Achilles of Rational Psychology*, eds. Tom Lennon and Robert Stainton (Dordrecht, the Netherlands: Springer Verlag 2007). For extensive discussion of the relevant issues about the notion of matter in the period, see Thomas Holden, *The Architecture of Matter* (Oxford: Clarendon Press, 2004).

noting that Clarke is remarkably quiet about what feature of consciousness requires what he calls an individual subject. In the Second Paralogism Kant explains the idea of the unity of consciousness employed in the argument as follows:

> For suppose it be the composite that thinks: then every part of it would be a part of the thought, and only all of them taken together would contain the whole thought. But this cannot consistently be maintained. For representations (for instance, the single words of a verse), distributed among different beings, never make up a whole thought (a verse), and it is therefore impossible that a thought should inhere in what is essentially composite. It is therefore possible only in a *single* substance, which, not being an aggregate of many, is absolutely simple.[16]

The idea of this argument is that the parts of a unified mental representation cannot be distributed over the parts of a composite subject, and for this reason the subject of thought must be simple. Clarke does state that consciousness cannot be the sum of a multitude of consciousnesses, but he never explains why this is so. And often the two interlocutors talk about the absurdity of the parts of a material subject of consciousness being conscious rather than about the problems for an aggregate consciousness.[17] Clarke identifies this claim with an admission that consciousness is not the sum of a multitude of consciousnesses.[18]

Collins agrees that consciousness cannot be the sum of consciousnesses of the parts of its subject, and so he does not probe Clarke on this claim. He does ask him what he thinks consciousness is and questions the idea that it requires an individual subject. But in response to such questions Clarke does not explicitly appeal to the kind of unity of consciousness considerations Kant discusses and that can be found in other early moderns.

Clarke does offer some clues about his conception of consciousness: when he offers the Achilles Argument, he distinguishes this argument, which focuses on "bare Sense or Consciousness it self[,]" from arguments that appeal to the higher capacities of the human mind: "its noble Faculties, Capacities and Improvements, its large Comprehension and Memory; its Judgement, Power of Reasoning, and Moral Faculties."[19] But what does Clarke mean by "consciousness?" He writes:

> *Consciousness*, in the most strict and exact Sense of the Word, signifies neither a *Capacity of Thinking*, nor yet *Actual Thinking*, but the *Reflex Act by which I know that I think, and that my Thoughts and Actions are my own and not Another's*. But in the present Question, the Reader needs not trouble himself with this Nicety of Distinction; but may understand it indifferently in all or any of these Significations; because the Argument proves universally, that Matter is neither capable of this *Reflex Act*, nor of the first *Direct Act*, nor of the *Capacity of Thinking* at all.[20]

---

[16] *Critique of Pure Reason*, A 352.

[17] See for instance Collins at W III.806.

[18] W III.798.

[19] W III.730.

[20] W III.784.

So Clarke offers a very specific characterization of consciousness as awareness that one's mental acts are one's own, but at the same time he does not think his argument focuses on this specific conception of consciousness; it concerns thinking in a very broad sense. Later he writes that he does not *need* to explain what consciousness is because "Every Man feels and knows by Experience what Consciousness is, better than any Man can explain it: Which is the Case of all simple Ideas."[21] And in the same vein he writes that we have "Intuitive Certainty" that consciousness cannot be a mode of motion.[22]

So Clarke does not offer us any real illumination on the question of what about consciousness requires an individual subject. Unfortunately, since Collins accepts that consciousness cannot be the sum of a multitude of consciousnesses, his probing on this issue is limited.[23]

2 Emergentism

Collins thinks that he can save materialism and avoid the problems raised by the Achilles Argument if the consciousness of a composite material subject results from other qualities that belong to the parts. In that case, the consciousness of the whole will not consist of a multitude of consciousnesses and the divisibility of matter poses no problem. In contemporary terms, one might say Collins proposes a type of emergentism. Using the term loosely, I will mean by emergentism the following: a configuration of qualities gives rise to a genuinely different kind of quality.[24] Contemporary philosophers distinguish a variety of forms of emergentism, but I will not attempt to try to identify just what type of emergentism Collins' proposal corresponds to. I will, however, address a striking ambiguity in the discussion of Collins' proposal.

---

[21] W III.790.

[22] W III.837. I discuss the question what Clarke means by consciousness at greater length in "The Achilles Argument and the Nature of Matter in the Clarke–Collins Correspondence." As I discuss there, this question is connected to the question what Clarke's precise ground is for the impossibility of consciousness belonging to a composite. In the tradition one can find two types of Achilles Argument: the type discussed by Kant which is based on the need to unify the contents of consciousness and a different type that relies on an analysis of self-consciousness. Clarke's definition of consciousness as "the *Reflex Act by which I know that I think, and that my Thoughts and Actions are my own and not Another's*" suggests this latter version, but his claim that this definition is not crucial renders that interpretation uncertain.

[23] Clarke and Collins do engage in an exchange about the question whether they are talking about actual or potential consciousness. I don't think this discussion adds anything substantial to the debate.

[24] For discussion of contemporary notions of emergentism, see Timothy O'Connor and Hong Yu Wong, "Emergent Properties," *The Stanford Encyclopedia of Philosophy* (Winter 2006 Edition), ed. Edward N. Zalta, <http://plato.stanford.edu/archives/win2006/entries/properties-emergent/>.

Collins introduces emergentism as follows. He contends that we frequently encounter examples of qualities or powers that belong to the parts of a complex material system and give rise to novel qualities or powers in the whole. As examples he offers the scent of a rose, the harmony produced by a musical instrument, the capacity of a clock to tell time, the development of sensation in a chick in an egg. His first example he describes as follows:

> And Matter of Fact is so plain and obvious, that a Man cannot turn his Eye but he will meet with Material Systems, wherein there are individual Powers, which are not in every one, nor in any one of the Particles that compose them when taken apart, and considered singly. Let us instance for example a *Rose*. *That* consists of several Particles which separately and singly want a Power to produce that agreeable Sensation we experience in them when united. And therefore either each of the Particles in that Union contributes to the Individual Power, which is the external Cause of our Sensation, or else God Almighty superadds the *Power* of producing that Sensation in us upon the Union of the Particles. And this, for ought I can see, may be the case of Matter's Thinking. Those Particles which compose the Brain, may under that Modification either have the Power of Thinking necessarily flowing from them, or else may have the Power of Thinking superadded to them by the Power of God, though singly and separately they may not have the Power of Thinking.[25]

So the parts of the rose taken by themselves lack the power to produce the sensation of scent in us, but, Collins proposes, either the power of the whole rose to produce the sensation of smell in us results from the qualities of the parts, or God superadds that power. The discussion focuses on the former possibility, and so will I. So Collins suggests that a genuinely new power or quality can result from a configuration of powers or qualities of the parts of a material system.

Clarke rejects emergentism using the following strategy. He introduces the Homogeneity Principle and a division of qualities into three kinds. The HP applies to the first type of quality, what Clarke calls "really inherent qualities"; consciousness is one of these. He argues that Collins' counterexamples are not really inherent qualities, but fall under the second or third type, and so they do not count against the HP. Clarke states the HP as follows:

> … [I]t is evident at first sight that every Power or Quality that is or can be *inherent* in any System of Matter is nothing else than the Sum or Aggregate of so many Powers or Qualities *of the same Kind*, inherent in all its Parts. The *Magnitude* of any Body is nothing but the Sum of the Magnitudes of all its Parts. Its *Motion*, is nothing but the Sum of the Motions of all its Parts. And if *Cogitation* in like Manner could possibly be a Quality really *inherent* in a System of Matter, it must likewise necessarily be the Sum and Result of the Cogitations of the several Parts: and so there would be as many distinct Consciousnesses as there are Particles of Matter, of which the System consists; which I suppose will be granted to be very absurd. Compositions or Divisions of Magnitude, varied in infinite Manners to Eternity, can produce nothing in the whole System no Quality or Power whatsoever but mere *Magnitude*; Compositions and Variations of *Motion*, nothing but *mere Motion*.[26]

[25] Collins, W III.751–752.

[26] Clarke, W III.759.

Obviously, crucial to the HP is the question when qualities count as being of the same kind. Clarke is aware of this problem and notes that "the Terms, Kind and Species, and of the same Kind or Species, are very ambiguous terms and used in great Variety of Significations."[27] The two correspondents discuss the issue at some length. Clarke explains that by qualities of the same kind in the HP he is not thinking of qualities of the same "*species specialissima*" but "*species generaliores*." For example, an instance of a specific type of shape must result from shapes, which may, however, be other species of shapes. They discuss the example of roundness: it is not the sum and result of roundnesses (it is not the case that "*Globosity* is made up of *Globosities*," Clarke writes) but "a *whole Round Figure* must necessarily be made up of *Pieces of Roundness*, which are all of the *same Kind* with it."[28] And a magnitude of a foot "is not an Aggregate of *Cubic Feet*, but of other *Magnitudes* which constitute a Cubic Foot."[29] Figure and motion, however, are not of the same species as thought at all, except in the sense that they are all qualities, Clarke notes. So they do not share a *species generalior* that is a subspecies of quality, and this is what Clarke requires for the HP.

This appeal to a classification of (really inherent) qualities raises further questions. Clarke's use of traditional Latin terminology means an implicit appeal to the tradition of classification under Aristotle's categories. But in the present context one might well want a defense of the classification of qualities into different types. Clarke does not offer anything of the sort. And we should probably not expect such a defense since, as we shall see later, he held that we know intuitively that consciousness and motion are different and have nothing in common, and this may manifest a broader view of his on the matter of the differences between qualities.

So the first component of Clarke's response to emergentism is the HP. In addition, he offers a three-part distinction of qualities or powers, similar to, but not identical with Locke's tripartite division in his discussion of secondary qualities in *Essay* II.VIII. Locke's division was limited to qualities of bodies, but Clarke's is not; it is intended as an entirely general classification of qualities. The first type of quality consists in really inherent qualities, which include consciousness, but also size and motion. To these the HP applies: configurations of size give rise to sizes, similarly for motions.[30] HP does not apply to the two other types of qualities, and Clarke argues that Collins' counterexamples all belong to the second and third types of qualities. Here is his description of these types:

---

[27] W III.827.

[28] W III.828.

[29] W III.828.

[30] W III.759. Clarke also uses examples of sounds, colour, smell, but it is problematic for him to do so given his mechanistic analysis of secondary qualities and given that he immediately classifies sweetness and colour as belonging to the next category of qualities (W III.759). Indeed, he then specifies that only as sensations in our thinking are they individual powers, but in the bodies they are at best "specifically, not individually, single powers; that is, they are only a number of similar motions or figures of the parts of the body." (W III.760)

> *Secondly*, Other Qualities there are, which are *vulgarly* looked upon as *Individual Powers*, resulting from and residing in the whole System, without residing particularly in each or any of its single and original Parts; such as are the *Sweetness* of certain Bodies, their *Colours*, etc. But this is only a *vulgar* and very *gross* Error. For neither do these Qualities reside in, or at all result from, the whole System, in any *proper Sense*: neither in *any Sense at all,* in which they can be ascribed to that Body or System of Matter to which they are vulgarly supposed to belong, are they truly *Individual Powers*. In the *first* place they are not *really Qualities* of the System, and evidently do not at all in any *proper Sense* belong to it, but are only *Effects* occasionally produced by it in some other Substance, and truly Qualities or Modes of that other Substance in which they are produced: thus the *Sweetness of a Rose*, is well known not to be a Quality really inhering in the Rose; but a *Sensation*, which is merely in him that smells it, and a Mode of the *Thinking Substance* that is in the Man …. And the same may be said of *Heat, Light, Taste, Sound* and all those others which we call *Sensible Qualities*. *Thirdly*, other Powers, such as *Magnetism*, and *Electrical Attractions*, are not *real Qualities* at all, residing in any Subject, but *merely abstract Names* to express the *Effects* of some determinate Motions of certain Streams of Matter; and *Gravitation* itself, is not a *Quality inhering in* Matter, or that can possibly *result* from any Texture of Composition of it; but only an *Effect* of the continual and regular Operation of some other Being upon it; by which the Parts are all made to tend one towards another.[31]

So Clarke divides our ordinary-life attribution of a quality to a substance into three types:

(i) Qualities that we attribute to a substance and that genuinely inhere in that substance. To these the HP applies.

(ii) Qualities that we attribute to a substance but that are really effects it produces in another substance. Secondary qualities belong to this category, such as the sweetness of a rose, which is really a sensation in us.

(iii) For the third type Clarke does not offer a clear definition. It seems like a fairly loosely defined category that applies when qualities don't belong to the first two types. What is clearly crucial for Clarke is that this category covers cases where we attribute qualities to a substance that are not really inhering qualities. We use "merely abstract names" for complex phenomena. He writes that this category comprises abstract names we use "to express the *Effects* of some determinate Motions of certain Streams of Matter [,]" but the category is broader than that. Gravity, for Clarke, is the result of an operation by God on matter. He also thinks that when we talk about a collection of qualities in a substance as if one thing we really have an abstract name at hand.[32]

---

[31] W III.759–760.

[32] For instance, "The Power of a *Clock* to show the Hour of the Day, is nothing but one *new complex Name*, to express at once the several Motions of the Parts, and particularly the *determine Velocity* of the last Wheel to turn round once in twelve Hours" (W III.797). One worry about Clarke's position is this: it seems important to the HP that qualities like motion abide by HP and so the motion that results from a combination of motions (and similarly for sizes, shapes) is a genuine inherent quality. But the notion of a really inherent quality suggests the idea that the quality is not a mere aggregate of qualities. Leibniz would object that Clarke cannot say this about qualities like motion, size and shape, and so their status as really inherent qualities seems unstable. This raises questions about the applicability of the HP generally.

The label "really inherent" for the first type is significant: Clarke is distinguishing attribution in a broad sense from genuine inherence. Our attributions of the other types of qualities do not reflect genuine inherence of a corresponding quality in the substance in question. Clarke's rejection of the example of the rose illustrates the point. In classical mechanistic fashion, he distinguishes between the sensation of scent in us and physical causes in the rose, and he analyzes the power in the rose to produce the scent in us in terms of a configuration of primary qualities, sizes and motions.[33] Indeed, he identifies this power with the collection of sizes and motions. There is then in his view no genuine new quality in the rose, just this collection of qualities of the parts.[34] In the case of the second and third type, we may make attributions of qualities that are quite different from the qualities of the parts that underlie them, but such attributions do not reflect genuinely inhering qualities.

So Clarke's strategy is to argue that Collins' counterexamples are not instances of really inherent qualities. Consciousness, however, is such a quality and since HP only applies to such qualities, the counterexamples are irrelevant. But why should one accept the Homogeneity Principle?

3 Clarke's Defense of the Homogeneity Principle

Consider the following statements in defense of the HP:

> Whatever can arise from, or be compounded of any Things; is still only those very Things, of which it was compounded …. For instance, All possible Changes of *Figure*, are still nothing but *Figure*; [All possible Variations, Compositions and Divisions of *Magnitude*, are still nothing but *Magnitude*;] … All possible Compositions or Effects of *Motion* are nothing but *mere Motion* … And how many other Qualities soever, *known* or *unknown*, the Particles of Matter be supposed to be indued with; those Qualities can never in any Composition or Division produce any new Power specifically different from themselves, unless a *Cause* could give more to the *Effect* than is in itself.[35]

> And this is evidently making a *Whole* bigger than *All its Parts*, that is, containing something different from, something over and above, something more than All its Parts

[33] There is some misunderstanding between the interlocutors about the example of the rose, as Clarke thinks that Collins was ascribing to the rose the scent as we experience it. Collins had not done so (W III.770).

[34] W III.790. Clarke sometimes classifies the power of the rose to produce scent in us as a quality of the second, sometimes as a quality of the third kind (W III.797). This is confusing and perhaps inconsistent, but I do not think it affects his argument, since what matters for Clarke is to rule out that various qualities are really inherent qualities.

[35] W III.788. Clarke is quoting here from his own *Demonstration of the Being and Attributes of God* (DBAG), Prop 8 section z, which can be found in the edition by Ezio Vailati (Cambridge: Cambridge University Press, 1998).

> taken together; nay such a Whole, the Sum of whose Parts neither make up the Whole itself, nor any Part of it: which is a plain Contradiction.[36]

Clarke's defense relies on two ways of conceiving the relationship between qualities of the whole and those of its parts: he conceives of the qualities of the whole as the *sum* of the qualities of the parts, and as their *effect*—without distinguishing between these two ideas. So he thinks of emergence both as the qualities of the parts constituting the quality of the whole and as causing the qualities of the whole. I will return to this point below. Furthermore, Clarke's defense of the HP contains two strands. The first relies on constraints on causality commonly accepted in the period, the latter turns on the notion of inherence. The concerns about inherence constitute the main objection against the possibility of superaddition by God. The causal constraints are central in Clarke's dismissal of emergentism, where, rather than being added by God, thinking results from a configuration of material qualities.[37] And so the causal constraints are particularly important to our concerns.

This type of causal constraint was widely accepted in the early modern period before Hume, but Clarke's use of them is atypical because he employs them to analyze the relationship of the qualities of a composite whole to those of its parts. Their use is better known for instances of causation of an effect in a subject that is different from the agent and of which the agent is not a part. A prominent place is Descartes's Third Meditation, where he wrote that the cause (the complete efficient cause, that is) must contain at least as much reality as the effect. Descartes focuses on levels of reality or perfection, and Clarke does the same in his *A Demonstration of the Being and Attributes of God* (DBAG), where he discusses the nature of God and argues for God's existence.[38] But when he

---

[36] W III.833.

[37] Clarke explicitly connects the concern about inherence to superaddition, when he writes that superaddition means "that a *Quality* is by the Power of God made so to arise out of Nothing as to be superadded to a *Subject*, and to subsist without inhering in that Subject, to which it is at the same time supposed to belong." (Clarke, W III.760; see also p. 759) I am separating the inherence and causal constraints sharply here, but perhaps more sharply than Clarke himself did. One can see the two constraints as connected by the following concern: the quality of the whole must be grounded in the qualities of the parts. The concern with inherence does not arise in the same way for emergentism as it does for superaddition, because emergentism proposes to ground the quality of the whole in the qualities of the parts insofar as they result from them.

[38] It is controversial whether Descartes' causal constraints require similarity between cause and effect or merely appropriate levels of reality. For discussion see Janet Broughton "Adequate Causes and Natural Change in Descartes's Philosophy," in *Human Nature and Natural Knowledge: Essays Presented to Marjorie Grene on the Occasion of Her Seventy-Fifth Birthday*, eds. Alan Donagan, Anthony N. Perovich Jr., and Michael V. Wedin (Dordrecht, the Netherlands: Reidel, 1986), pp. 107–127; Eileen O'Neill (1987), "Mind–Body Interaction and Metaphysical Consistency: A Defense of Descartes," *Journal of the History of Philosophy* 25: 227–245; Margaret Wilson (1991), "Descartes on the Origin of Sensation," *Philosophical Topics* 19: 293–323; Tad Schmaltz, "Sensation, Occasionalism, and Descartes' Causal Principles" in *Minds, Ideas and Objects: Essays on the Theory of Representation in Modern Philosophy*, eds. Philip D. Cummins and Guenther Zoeller (Atascadero, CA: Ridgeview, 1992), pp. 38–55;

discusses the issue of emergentism in the correspondence with Collins, Clarke does not talk about levels of reality or perfection. Rather he focuses on the question whether qualities are sufficiently *similar*. He often illustrates the HP by saying that compositions of motion or magnitude will only result in more motions or magnitudes and cannot produce qualities of a different kind.[39]

What is the intuitive appeal of these causal constraints? The model of causation at work is of course intensely pre-Humean: causation is not merely a matter of constant conjunction, and it is also not simply a matter of causal laws. The model is made intuitive by examples like heat (an example Descartes uses) or motion: one ball makes another ball move, but the first ball must have at least as much motion (or at least as much force, Leibniz would say) as the second ball. One body heats up another one, but it can't produce more heat in the second body than it contains in itself. If that were to happen some of the heat would come from nothing. It is a model that relies on genuine causal agency, and the idea that in causation some entity is produced. The model suggests some stuff flows from the agent to the patient, or is passed on from one to the other. The transmission model finds clear expression in late scholasticism. So we find Suarez writing that causation "is nothing other than that influx or concourse by which each cause in its kind actually flows [*influit esse*] being into the effect."[40] Clarke speaks of the cause giving to the effect.[41] Causal constraints like the HP embody the idea that a quality that is produced is an entity that must come from somewhere, and all of it must come from somewhere. Otherwise it, or some of it, comes from nowhere, which is impossible. Crucial to Clarke's use of the constraints is that he sees the constraints not as merely quantitative; his focus is on qualities that are different in

Schmaltz, "Causation and Similarity in Descartes" in *New Essays on the Rationalists*, eds. Rocco J. Gennaro and Charles Huenemann (Oxford: Oxford University Press, 1999).

[39] While the application of the causal constraint on causation within a substance is less well known, it also lurks in the background in Locke's argument that God must be an immaterial substance because the cause of thinking beings like us must be itself a thinking being. Locke argues that matter itself cannot produce motion *in itself*, and similarly "… Matter, incogitative Matter and Motion, whatever changes it might produce of Figure and Bulk, *could never produce Thought*: Knowledge will still be as far beyond the Power of Motion and Matter to produce, as Matter is beyond the Power of nothing, or nonentity to produce." (*Essay*, IV.x.10, p. 623) The idea that inert matter could produce motion in itself, or that matter in motion could produce thought is ruled out because it would be like something coming from nothing, just as is the case for Clarke for material qualities producing consciousness. It is not always clear in this section of the *Essay* whether Locke is talking about causation within the same subject or in another subject. In fact he seems to move freely between the two.

[40] *Disputationes metaphysicae*, 2 vols. (Georg Olms Verlag, 1998), XII.ii.13. And Eustachius of St Paul: the formal definition [*ratio*] of causation "is placed in a real influx of the cause into the effect: so that to cause an effect is nothing other than to really flow into this effect by communicating being to it." (*Summa philosophica quadripartita* (Paris: Carolus Chastellain, 1609), III 52.

[41] W III.788. Jonathan Bennett discusses the model as a kind of giving. See his *Learning from Six Philosophers* (Oxford: Clarendon Press, 2001), vol. 1, pp. 84–86.

kind: if the purported resulting quality is qualitatively radically different from the purported originating qualities, it comes from nowhere.

There are various aspects of Clarke's use of causal constraints that may strike one as puzzling. First, as I noted, a difference between Clarke's use of the HP in rejecting emergentism and Descartes's better-known use of causal principles is that the latter is talking about causation of an effect by a subject in a subject distinct from the agent of which the agent is not a part. The examples of heat and motion make qualitative and quantitative constraints on causation intuitive for this type of causation. But Clarke talks about the production of a quality of a composite by qualities of the parts of that same composite. Furthermore, he often talks as if obedience to the HP is a matter of simple addition rather than causation, and he does not separate the two ideas. In other words, he seems to conflate two models of emergence: one sees the emerging quality as constituted by the underlying qualities, the other as caused by them.

These would seem to be two very different models of emergence. But the two are not so far apart given the model of causation as a kind of transmission of something: if the causes are the qualities of the parts and the effects the qualities of the whole, one can see how the qualities of the parts can be seen as adding up to qualities of the whole, and one can perhaps also see them as giving to the whole in a broad sense so that the magnitude of the whole results.[42] Furthermore, one can combine the two seemingly different ideas into one by means of what one might call "constitute causation", a type of causation that specifically covers the explanation of the quality of a composite in terms of the qualities of its parts.[43]

---

[42] Clarke's favourite examples of obedience to HP are motion and magnitude, and he offers his most detailed analysis for the case of roundness: the roundness of a body results from the convexity of its parts. This is easily understood as a case of wholes and parts being added up. The examples are instances of traditional mechanistic qualities, motion and magnitude, which are quantifiable and can easily be added up. So for the examples Clarke has in mind, it is easy to understand his tendency to talk about the HP in terms of addition and to present it as a quantitative constraint. And it is fairly easy to see why one would identify the causal and constitutive understandings of emergence. But how about other types of qualities?

Clarke does discuss the application of HP to the mixing of colours: "When the Mixture of Blue and Yellow Powder makes a Green, that Green is still nothing but Blue and Yellow intermixt, as is plainly visible by the Help of Microscopes." (W III.788) But given that he sees a colour insofar as it can be attributed to a body as a configuration of mechanistic qualities, this application of HP to colour is just an application to mechanistic qualities again. He does not discuss how we should understand sensible qualities insofar as they are sensations in the mind: does the HP apply to our sensations of colour? It makes sense to think that colours can only be mixed with colours to get a new colour, as opposed to a sound or flavour. But it is not clear how this would work with sensations, and the quantitative aspect of the HP is hard to apply to sensations. So while the HP is formulated in general terms about any type of quality, it in fact seems more clearly applicable to mechanistic qualities primarily than to other types of qualities.

[43] For discussion of a version of this type of causation, see John Searle, *Minds, Brains, and Science* (Cambridge, MA: Harvard University Press, 1984), pp. 19–23. I am grateful to Daniel Warren for this point. Although Clarke himself does not clearly formulate such a notion of causation it nicely captures his view. One question is how this notion exactly relates to efficient causation. Clarke himself uses the

A different concern arises from the fact that Clarke was a dualist who believed in mind–body interaction. So one might well ask: what about the production of mental states by physical states and vice versa? Doesn't this violate the causal constraints given the dissimilarity between the two? Collins does query Clarke about interaction, which he rightly points out is a difficult issue for dualism. Clarke does have a way of reconciling HP and mind–body interaction, however. He writes that bodies can produce sensations only because the power of thinking already exists in the mind:

> For the Power that is in one Substance, of exciting different Modes in another Substance; presupposes necessarily in that other Substance the *Foundation* of those Modes. Thus in the Case of all the *sensible Qualities* of Bodies; the *Power of Thinking* is beforehand in that Being, wherein those Qualities excite or occasion *different Modes* of Thinking.[44]

So it is not the case that a bodily process produces by itself a novel quality in the mind that is quite different from the material qualities that are its cause. That would be a violation of the causation constraints and the resulting sensation would come from nothing, its occurrence would not be (fully) explained. Instead the bodily quality results in the sensation in the mind because the mind contributes to the production of the sensation in virtue of having a power of thinking already in it. In other words, the full explanation of the occurrence of the sensation appeals both to the action the body and to the nature of the mind, the subject of the resulting state. So for Clarke the sensory state does not come from the bodily state in a way that requires that the bodily state be like the sensation in the sense required by the HP.[45]

While the HP and similar early modern constraints on causation have intuitive force given a certain model of causation, it is worth contemplating just how strict these constraints on causality are from a historical perspective. These constraints do not only distinguish many early moderns from Hume, but also from the Aristotelians. The early modern mechanists thought of all physical processes as rearrangements of particles of a homogenous matter, and for them this was a great virtue as it meant that their picture of the physical world was marked by a high level of simplicity and intelligibility. They repeatedly claimed that their model of causation in the physical world was more intelligible than that of the

---

language of efficient causation, for instance, he states that a violation of HP would mean that "something would without any Efficient, be produced out of Nothing" (W III.786).

[44] W III.797.

[45] Clarke's use of the term "occasion" does not mean that he was an occasionalist; he was not. For discussion see Vailati (1997, pp. 58–59). Clarke may have intended to express a kind of model Steven Nadler has labeled "occasional causation." This is a complex causal model where both body and mind play a role. See Nadler (1994b), "Occasional Causation," *British Journal for the History of Philosophy*, 2: 35–54. I discuss Descartes's use of this kind of causal model for sensation in my "Descartes on Mind–Body Interaction: What's the Problem?" *Journal of the History of Philosophy* 37: 435–467. My conception of this causal model for Descartes deviates from Nadler's in that I think for Descartes at least (as opposed to perhaps other early moderns) the body does act as a genuine *efficient* cause on this model in sensation.

scholastics.[46] Thus Descartes compared his own mechanism to the scholastics as follows:

> We understand very well in what way the various local motions of one body are brought about by the different size, shape and motion of the particles of another body; but we can not at all understand in what way those very same things (namely size, shape and motion) can produce something else that is entirely different from them in nature, as are those substantial forms and real qualities, which many suppose to be in things; nor in what way those qualities or forms then have the power to bring about local motions in other bodies.[47]

In his treatise *The World* Descartes offers the example of a fire: he understand the process in terms of mechanistic qualities, and criticizes the Aristotelians who admit interaction between mechanistic qualities and secondary qualities like heat and colour realistically understood.[48] So compared to the early modern mechanistic picture, the Aristotelian view of the world presents a more hybrid and more opaque array of causal processes.

Particularly interesting in relation to our concern with emergentism is the scholastic conception of the foundations of various kinds of qualities in the physical world. They saw the elements as fundamental, and these were characterized by what they called *primae qualitates*, hot and cold, dry and wet. Other qualities, commonly labeled *qualitates secundae*, which included tactile qualities and nontactile sensible qualities, like colours, smells, arise from the *primae qualitates*, but in ways that were not really made clear.[49] This is a kind of emergence that would be ruled out by Clarke's HP. So the similarity constraints on causation in Clarke and other early moderns do not merely contrast with Humean causation, but also with the earlier Aristotelian conceptions. By contrast with the later Humean picture, the early modern mechanists explicitly aimed to provide causal models that really explain and make intelligible the occurrence of an effect in light of the cause, and they thought their model was superior in this regard to the Aristotelian model.

One way to reject Clarke's position is to reject his qualitative constraints on causation. One could do this by becoming a Humean about causation. Collins does

---

[46]I discuss this issue for Descartes in my *Descartes's Dualism* (Cambridge, MA: Harvard University Press 1998), pp. 135–137; for Locke in Rozemond and Yaffe (2004) in "Peach Trees, Gravity and God: Locke on Mechanism,", *British Journal for the History of Philosophy*, 12: 387–412.

[47] *Principles* IV.198.

[48] AT XI.7–10, CSM I.83–84.

[49]For discussion, see Anneliese Maier, "The Theory of the Elements and the Problem of their Participation in Compounds," in *On the Threshold of Exact Science*, ed./transl. Seven Sargent (Philadelphia, PA: University of Pennsylvania Press, 1982), pp. 124–142, esp. pp. 135–139. At the same time, causal likeness principles do play a role in scholastic accounts of causation. See, for instance, Aquinas, *Questions on the Soul*, qu. 12 James Ross, transl. (Milwaukee, WI: Marquette University Press, 1984), p. 157. I can't do justice to the complexity of the issues in Aristotelian scholasticism here.

not do this, and the world had to wait another while for Hume. Instead, remarkably, in his third letter Collins writes that he does accept the HP—if understood properly. He distinguishes between generical and numerical powers:

> By *Numerical Powers* I understand such Powers as Motions and Figures of the same Species. The Power of the Eye to contribute towards seeing, is a Species of Motion, and the Roundness of a Body is a Species of Figure. By *Generical Powers* I understand all the several Species of *Numerical Powers*; as Motion signifies all the various Species of Motion, and Figure all the various Species of Figure.[50]

He then claims he does accept HP if applied to generical powers, but not if applied to numerical powers:[51] the shape of a body is the sum of the shapes of the parts, but the roundness of a body is not the sum of the roundnesses of the parts.

Clarke rejects Collins' distinction between generical and numerical powers on the ground that generical powers are universals and in things powers are always numerical.[52] But Collins had said so himself,[53] and in fact the point Collins is trying to make is what Clarke himself has in mind when he writes that "qualities of the same kind" refers to qualities that belong to the same "*species generalior*." So at this point they seem to agree on the following version of the HP: the quality of a composite subject must be the sum and result of qualities of its parts that belong to the same *species generalior* as the quality of the whole they constitute.

Collins has not suddenly become a dualist, however. He makes a new materialist proposal: he proposes for consideration the possibility that consciousness might be a modification of motion in the way in which roundness is a mode of figure; that is to say, it is a specific type of motion. And he proposes that this particular type of motion might result from a combination of motions of the parts in accordance with the HP. Collins insists that his point is not to claim that consciousness really is a type of motion. Rather his point is to consider the view that consciousness is a mode of some power or other of matter, but for the sake of argument he considers the specific possibility that consciousness is a mode of motion.[54]

Collins' acceptance of HP and his shift to the proposal that consciousness be identical with a mode of motion looks like a significant change of heart on Collins' part. Why, one might ask, this change? There is no clear answer to this question in the texts, but it is not clear that it is a real change. His acceptance of the HP comes as a result of substantial clarification of the principle. So it is possible that in its specified form Collins accepted the principle all along and that consciousness being a type of motion is what he had in mind from the beginning.

This move does substantially shift the debate from the question of any sort of robust emergentism to a proposal of identity of consciousness with a material

[50] W III.805.

[51] W III.806.

[52] W III.829.

[53] W III.806.

[54] W III.806, 859.

quality, and Clarke responds accordingly. And so in the end for both Clarke and Collins the HP rules out emergentism: a configuration of qualities can give rise only to qualities of a complex whole that are suitably similar, and so no genuinely novel qualities can result from a configuration of qualities of the parts of a complex subject.

4 The Limits of Our Knowledge

Collins' change of course comes with an element of constancy, however: in true Lockean fashion, throughout the exchange he appeals to the limits of our knowledge. Crucial to Locke's claim that we cannot rule out the possibility that God superadds thinking to matter was his view that we do not know enough about the nature of thinking or material substance to rule out this possibility. Collins goes beyond Locke when he suggests that thinking might result from material qualities. In doing so Collins repeatedly claims that Clarke is too optimistic about the scope of our knowledge:

(1) He claims that Clarke fails to show that his tripartite division of qualities is exhaustive. There may be really inherent qualities that do not obey HP.[55] The result of this part of the discussion is a disagreement about the burden of proof: should Clarke show his division is exhaustive, or should Collins show it is not?[56] And Clarke responds to this objection that he has *argued* that the HP applies to really inhering qualities. Furthermore, we saw that in the end Collins grants the HP. So the role of this claim of modesty about our knowledge is superseded by later developments in the discussion.

(2) Clarke presupposes a standard list of primary qualities for matter: he discusses size, shape and motion. But sometimes Collins suggests that there may be types of material qualities we are ignorant of.[57] Collins does not elaborate on this suggestion as much as he might. But in light of the changing conceptions of matter over the course of the centuries, this surely is a point that has significant force. I will return to it below.

(3) Collins suggests that we do not know enough about consciousness or motion to rule out the possibility that consciousness is a mode of motion. He writes: "… Consciousness, of whose Nature we are ignorant, may inhere in a System of Matter, without being the Sum of the Consciousnesses of the Parts."[58] He also thinks Clarke is too optimistic about our knowledge of motion. He writes that Clarke had failed to consider particular modes of motion, and that we do not

[55] W III.767.

[56] W III.803–804.

[57] W III.803–804, 806.

[58] W III.806.

have ideas of all the types of motion.[59] This is not the place for a full treatment of Clarke and Collins's discussion of the possibility that consciousness is a type of motion, but I wish to explore briefly this discussion insofar as it concerns the issue of our knowledge of motion and consciousness.

As I noted before, Clarke thought we all know what consciousness is, and he thinks the ideas of consciousness and motion are simple ideas we can't explain. He thinks that while there may be limits to our knowledge of motion and consciousness, our ideas of motion and consciousness are clear enough to rule out that consciousness is a type or mode of motion. He appeals to his claim that consciousness and motion have no common genus. He concludes from this that we have intuitive certainty that consciousness is not a mode of motion, just as we know that "a *Circle* or a *Cube* is not a *Thought*, or that an *Acute Sound* is not a *Purple Colour*."[60] He also specifically addresses Collins' claim that consciousness could be a specific mode of motion:

> Every *Mode* of any *Power* or *Quality*, is nothing else but *That Power* or *Quality*, of which it is a *Mode*, understood with some particular Limitation; that is to say, it is nothing but a *particular Instance* of that *general Power* or *Quality*; nothing but the *general Power* or *Quality*, considered under this or that *particular Modification*. *Blue* and *Red*, and all other *Modes* of *Colour*, are nothing but several particular *Colours*; and can contain nothing in their Idea beyond the *Genus* of *Colour* …. Now if *simple Ideas* be the Foundation of our Knowledge; and *clear and distinct Perception* of the *Agreement* or *Disagreement* of those Ideas, be the best and greatest *Criterion* of Truth, that our Faculties inable us to attain to; then it is as evident as any Truth in the World, that *Consciousness* cannot possibly be a *Mode of Motion*. For I have as *clear and distinct* a *Perception*, that the Idea of *Consciousness* contains something in it besides and beyond the *Genus of Motion*, as I have that it contains something in it beyond the *Genus of Figure*.[61]

Clarke surely has a point in saying that specific types of motion will be variations of the general kind and that at the same time consciousness does not seem to be such a variation. The possibility that consciousness might be a type of motion strikes me as not particularly promising, it seems hard to make sense of this identity. Clarke thinks the same goes for figure, and this too seems plausible: it strikes me as hard to make sense of such an identity.

Collins does not agree with this analysis. Whose side one takes on this issue might depend on whether one minds an identity between items where the identity does not seem at all intelligible. If causal connections can be brute and come without providing an understanding why or how A causes B, similarly one might think an identity could be brute. We might have no understanding how A and B could be identical, but we might not regard this as an obstacle to identity. Indeed, in light of this consideration it is striking that Collins proposes an identity of consciousness and motion, qualities that surely are intuitively quite dissimilar, while granting the HP, which rules out that consciousness emerges from motions.

---

[59] W III.806.

[60] W III.837.

[61] W III.836–837.

Collins was unimpressed by Clarke's arguments to the effect that we know enough to rule out an identity between consciousness and motion.[62] But even if we accept Clarke's claim that thinking cannot be motion or figure, we may refuse to follow him all the way. Clarke makes a much stronger claim: he thinks that the arguments that show that thinking cannot be a mode of motion also show that "it is not possible for *Thinking* to be a *Mode of Figure*, or of any other *known* Property of Matter*:* And also that it is not possible for it to be a *Mode* of any *unknown* Power of Matter, which in the general is *void of Thinking*."[63] So he thinks he can rule out that consciousness can be a mode of any type of material quality.

I do not see what justifies this optimism.[64] Clarke is right to note that one does not necessarily need to know everything about a quality in order to rule out its identity with another quality. But the broad claim that he has refuted thinking being a mode of any kind of material quality requires more than refuting its identity with some particular types of qualities. Clarke would either need to show that he has a full list of the qualities of matter, or, alternatively, he could argue that something about the nature of matter in general means that whatever qualities it might turn out to have cannot be candidates for identity with thinking.

The Achilles Argument is such an argument: it contends that the nature of matter is such that it cannot be the subject of thinking because matter is divisible and consciousness cannot be composite in the way in which qualities of matter must be. Collins's emergentism was meant to get around this by arguing that consciousness might emerge from material qualities and thus need not be the sum of the consciousnesses of the parts of a material subject. This approach invites a discussion of the relationship between consciousness and particular types of material qualities. But now what Clarke seems to need is either confidence that we know all the particular types of material qualities or an argument that restricts what kinds of material qualities are possible. The latter seems like a more promising approach. The Achilles Argument is successful in a context where assumptions about the nature of matter as inherently composite dictate what sorts of qualities it can have. But on either of these approaches an argument for dualism risks being time-bound as a result of the changing conceptions of matter and its qualities. Over the history of philosophy and science conceptions of matter have changed radically. No doubt philosophers like Descartes or Clarke would be bewildered by current scientific conceptions of matter. Indeed, in Clarke's own day the question of the nature of gravity troubled the waters on the nature of

---

[62] W III.865–870.

[63] W III.836.

[64] Clarke's optimism is no doubt due in part to the fact that he thinks that Collins chooses to discuss the possibility that thinking is a mode of motion as an example, because motion is the most plausible candidate for identity with consciousness. And so although Collins insists that this is just an example of the sort of thing he has in mind, Clarke thinks it is legitimate to focus on the plausibility of that particular (W III.836).

matter, and this question is among the ones Clarke and Collins debate. One question is whether there is enough constancy in the conceptions of matter so that some form of the Achilles Argument could survive.[65]

Conclusion

The Clarke–Collins correspondence stands out in the early modern period for its detailed discussion of the possibility that consciousness emerges from material qualities. An examination of the exchange between Clarke and Collins reveals various significant ideas about how emergentism fits into an early modern context. The correspondence begins with Clarke offering the Achilles Argument, an argument that enjoyed considerable popularity in the early modern period. In Clarke's version, the argument contends that matter is divisible, but consciousness must belong to a simple subject because it cannot be the sum of a multitude of consciousnesses. So consciousness cannot belong to a material subject. Collins accepts the impossibility of consciousness being the sum of a multitude of consciousnesses but suggests that the problem is avoided if consciousness emerges from a configuration of material qualities of the parts. This part of the exchange points up a significant issue for the success of the Achilles Argument in establishing the immateriality of the mind: it assumes the impossibility of emergence.[66]

Clarke's Homogeneity Principle is central to the discussion: a quality can only result from qualities that are like it. In the end, the qualitative constraint on causality embodied in the HP rules out any substantive form of emergentism for both correspondents. During this period the constraints on causality were particularly strict compared both to previous Aristotelian conceptions as well as later Humean conceptions. Given the widespread acceptance of this kind of constraint by early moderns, its role in the exchange between Clarke and Collins suggests that for other early moderns also such constraints are likely to pose a serious obstacle to emergentism. Clarke's HP cannot be separated from his notion of an inherent quality: not just any attribution of a quality to a substance corresponds to that substance having a really inherent quality. Many of our attributions are abstract names for complex phenomena consisting of a collection of qualities within the substance in question or even belonging to several substances. HP applies to really inherent qualities and consciousness is one of these. So many purported counter-examples, according to Clarke, miss their target.

---

[65]A kindred line of thinking in contemporary philosophy is the idea that material explanations are structural in a way that consciousness is not. See David Chalmers, "Consciousness and its Place in Nature" in *Blackwell Guide to Philosophy of Mind*, Stephen Stich and Fritz Warfield, eds. (Oxford: Blackwell, 2003).

[66]As Karl Ameriks points out in his discussion of Kant's treatment of the argument. *Kant's Theory of Mind*, 2nd ed. (Oxford: Oxford University Press, 2000), pp. 58–59.

Finally, a consistent current of disagreement lies in Clarke's and Collins' differing views about the limits of our knowledge. This disagreement is particularly significant in their assessments of the possibility that consciousness might be a mode of a particular type of material quality. Clarke is optimistic that we know enough to rule out the identity of consciousness with a mode of motion, or any other kind of material quality. Collins' Lockean modesty about the scope of our knowledge leads him to say that there may be qualities of matter that we do not know about, and it leads him to go beyond Locke in denying that we can rule out an identity of consciousness with material qualities, his example being motion. This type of disagreement about the scope of our knowledge and what it means for what we can show about the relation between the mental and the physical is an important one in the early modern period. It clearly separates, for instance, Locke from Descartes, but also from Leibniz.

This examination of the disagreement between Clarke and Collins on the possibility of thinking matter is far from complete. But I hope that the reader will be convinced that the correspondence is rich in philosophically interesting considerations about this issue that preoccupied many thinkers in the early modern period. Furthermore, it offers insight into several important and interesting lines of thought that were shared by various philosophers in this period, and deserves much further investigation. Finally, there are significant resonances with current debates about the mind–body problem, which I have not had the opportunity to explore.[67]

[67] This paper has benefited from presentation at conferences at Oxford University, the University of Colorado at Boulder, Queen's University as well as to the philosophy departments at the University of Western Ontario, Kansas State University and Berkeley. In addition I am grateful for helpful discussions with Jessica Wilson, Donald Ainslie, and Lisa Shapiro, but especially Daniel Warren.

Berkeley and Hume on Self and Self-Consciousness

Talia Mae Bettcher

> For my part, when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never catch *myself* at any time without a perception, and never can observe any thing but the perception.[1]

So complained David Hume, who then went on to offer his conception of the self as an evolving bundle of perceptions, and who yet later in the Appendix, declared that his position was beset with difficulties.

In this paper I argue for a fresh understanding of Hume's famous Complaint (as well as his remarks in the Appendix). Specifically, I place the Complaint within the context of a dispute between Berkeley and Hume concerning self and self-consciousness.[2] I use Locke's innovative notion of *self* as the backdrop for my discussion.[3] My aim is to show the value of examining this important philosophical moment other than as a response to Descartes (as our prevailing cultural account would have it).[4] Instead, by looking to Berkeley we can secure a far deeper appreciation of what is at stake for Hume.

Before proceeding, let me make plain that I assume Hume read Berkeley. While this has been contested in the literature, it has now been established beyond doubt.[5] I assume Hume took Berkeley seriously, was influenced by him in important

---

[1] T I.iv.6.3,164; SBN 252. Citations of *A Treatise of Human Nature* [T] refer to book, part, section, paragraph, and page. All references are from David Fate Norton and Mary J. Norton, eds. (Oxford: Oxford University Press, 2000). I include page references for ed. L.A. Selby-Bigge with revisions and notes by P.H. Nidditch (2nd ed.) [SBN] (Oxford: Clarendon Press, 1978). Citations of *An Enquiry concerning Human Understanding* [EHU] refer to section, paragraph, and page. All references are from Tom L. Beauchamp, ed. (Oxford: Oxford University Press, 1999).

[2] Citations of *A Treatise concerning the Principles of Human Knowledge* [PHK] refer to part and section; of *Three Dialogues between Hylas and Philonous* [3D] to dialogues number and page; of *Philosophical Commentaries* [PC] to entry number; of *De Motu* [DM] to section and page, and of *Alciphron or the Minute Philosopher* [ALC] to dialogue number, section, and page. All references to Berkeley are from A.A. Luce and T.E. Jessop, eds. *The Works of George Berkeley, Bishop of Cloyne*, 9 vols. (London/Edinburgh: Thomas Nelson & Sons, 1948–1957).

[3] Citations of *An Essay concerning Human Understanding* [*E*] refer to book, part, section, and page from Peter H. Nidditch, ed. (Oxford: Clarendon Press, 1975).

[4] See Kim Atkins, ed. and commentary, *Self and Subjectivity* (Oxford: Blackwell, 2005), pp. 37–38.

[5] The question was raised by Popkin (1959) in "Did Hume Ever Read Berkeley?" *Journal of Philosophy* 56. Subsequently, a letter (August 31, 1737) was discovered in which Hume encourages Michael Ramsay to read Berkeley's *Principles* in order to better understand "the metaphysical parts of my reasoning" which was reprinted in Popkin's (1964) "So, Hume did

J. Miller (ed.), *Topics in Early Modern Philosophy of Mind*, Studies in the History of Philosophy of Mind 9, 193–222.

ways, and was concerned to address his views. Aside from internal evidence which indicates Hume's highest regard for Berkeley,[6] commentators have shown the influence on issues including *minima*, the contrast between the vulgar and the philosophical, the primary/secondary distinction, and even the mind itself.[7]

In the view I shall defend, both Berkeley and Hume are led to a transformed conception of self-consciousness owing to a shift away from the older substance-mode ontology. At its most detailed, the dispute concerns how best to conceptualize self-consciousness in light of this rejection of the older ontology.

For Berkeley, consciousness of one's own present existence is considered radically distinct from one's awareness of ideas. They are two discrete modalities of awareness, and the latter provides no interesting information about the intrinsic features of spirits. Moreover, because ideas are not viewed as intrinsic states of spirits, the "I" that is given in consciousness needn't be viewed as fleeting with each passing thought. Instead, spirits and ideas possess entirely different relations to time.

By contrast, for Hume consciousness of oneself is nothing other than consciousness of the various perceptions which together constitute the mind. Consciousness of oneself is always consciousness of one's *past* existence yielded through the associative mechanisms of memory. The very consciousness of one's own present momentary existence which serves to ground Berkeley's conception of spirit is left in a troubling place. I argue that Hume's famous remarks in the Appendix, far from indicating a serious problem with Hume's account, may constitute his own considered response to Berkeley's conception of self-consciousness.

While the details of this dispute are subtle, the stakes involved are philosophically large. Berkeley's own account of self-consciousness yields a view of the universe and our knowledge thereof which is fundamentally bifurcated into spirits and ideas. This, for Berkeley, leads to the view that while spirits are available to consciousness, they cannot be "objects of understanding"—they cannot be subject to scientific inquiry. For Hume, of course, the goal is to show

---

Read Berkeley," *Journal of Philosophy* 61. The discussion, however, still persisted. Michael Morrisroe Jr. (1973) provides an overview of the debate in "Did Hume Read Berkeley? A Conclusive Answer," *Philological Quarterly* 52 (2): 310–315.

[6] The internal evidence includes the reference to Berkeley as "a great philosopher" in Hume's *Treatise* discussion of abstract ideas (T I.i.7.1, 17, SBN 17), the *Enquiry* reference to Berkeley as a "very ingenious author" (EHU XII.v.203), and the reference to Berkeley's *Alciphron* in the 1763 essay "Of National Characters" in *David Hume: Essays, Moral, Political, and Literary*, ed. Eugene Miller (Indianapolis, IN: Liberty Classics, 1985), p. 209.

[7] Raynor (1980), "'*Minima Sensibilia'* in Berkeley and Hume," *Dialogue* 19: 196–199; M. R. Ayers, "Berkeley and Hume: A question of influence" in *Philosophy in History*, eds. R. Rorty, Schneewind, and Skinner (Cambridge: Cambridge University Press, 1983), pp. 303–327; D.R. Raynor, "Hume and Berkeley's *Three Dialogues*" in *Studies in the Philosophy of the Scottish Enlightenment*, ed. M.A. Stewart (Oxford: Oxford University Press, 1990), pp. 231–250. Especially noteworthy, Raynor (pp. 236–237) argues that Hume's account of mind may very well have been inspired by Berkeley's 1734 additions to the *Three Dialogues* (in particular, Hylas' proposal that the mind is a system of ideas).

how the mind *can* be subsumed into scientific investigation in a way that mirrors Newton's treatment of the world. Thus he has a strong interest in overturning Berkeley's account.

1 Locke on the Self and Self-Consciousness

Locke introduces 'self' as a quasi-technical expression. He characterizes it as "… that conscious thinking thing, (whatever Substance, made up of whether Spiritual, or Material, Simple, or Compounded, it matters not) which is sensible, or conscious of Pleasure and Pain, capable of Happiness or Misery, and so is concern'd for it *self*, as far as that consciousness extends."[8] This use of the term relates the self to consciousness, ontology, and self-concern. For the purposes of this paper, however, I put the last of these to the side. Suffice it to say that self-concern is determined by the limits of the self, which is in turn limited by the scope of self-consciousness.[9] Pain and pleasure play a special (limiting) role with regard to our concern,[10] and are the basis for human passion.[11]

For Locke, the self is related to consciousness precisely because consciousness *makes* an entity a self: "… 'tis that, that makes every one to be, what he calls *self*; and thereby distinguishes himself from all other thinking things …."[12] Although Locke does not say this, I take it that, more specifically, *self*-consciousness makes a being a self.[13] By contrast, the notion of self often relates to *ontology* in merely negative ways. For example, much of the chapter "Identity and Diversity" shows how the notions of *self* and *substance* come apart. Moreover, Locke's appeal to a self is best understood in terms of his claim that we do not know the real essence of the soul. Because we do not understand the essence of the soul, we do not know whether the soul is immaterial.[14] Such ignorance may cause problems with respect to our knowledge of the immortality of the soul. However, for Locke it is sufficient we have faith in a Judgement Day resurrection.[15] His theory of the self can render plausible this assurance by yielding an account of personal identity where present consciousness of one's past actions is sufficient to determine

[8] E II.xxvii.17, 341.

[9] E II.xxvii.17, 341.

[10] E IV.ii.14, 537.

[11] E II.xx.3, 229.

[12] E II.xxvii.9, 335.

[13] Because Locke characterizes 'person' in a way that requires both reason and reflection, I think 'self' and 'person' are not interchangeable. Rather, Locke appeals to the notion 'self' in order to explain how identity of the person can be secured.

[14] E IV.iii.7, 540–542.

[15] *Mr. Locke's Reply to the Bishop of Worcester's Answer to his second Letter* (1699) in *The Works of John Locke* 10 vols. (London: Thomas Tegg, 1823), vol. 4, p. 476.

identity over time. Locke's notion of a 'self' emerges as a modern substitute for the older notions of soul and substance; it ought to be understood as standing in a contrast with items it supersedes.

In order to press for this distinction between self and substance, Locke argues early in the *Essay* that Descartes is wrong to suppose that the essence of the mind consists in thinking. For Descartes accepts as a consequence of his claim that the essence of the mind is thought, the position that the mind always thinks; he denies that the mind remembers many of the thoughts that occur when sleeping.[16] And early in Book II of the *Essay* Locke attacks the view that the soul always thinks, and thereby the view that thought constitutes its essence.[17] Here Locke recognizes the supposition that the soul always thinks while we sleep to be counter-intuitive and worries that our failing to remember such nightly thought would lead to two persons. The real occurrence of frequent interruptions in conscious thought is therefore a central fact used by Locke to undermine the Cartesian account of mind.

Given Locke's recognition of this fact, it seems he is prepared to admit the intermittent existence of a self. During periods of sound sleep, there is no consciousness to constitute the self at that moment; and since there is no consciousness at that moment there is generally nothing to be remembered later on. Since the moment of sound sleep will generally not be appropriated to this self through what Locke calls its "consciousness of past actions," it will not be a moment at which the self can be said to exist.

Unfortunately, this is in *prima facie* tension with Locke's claim that one thing cannot have two beginnings,[18] and it raises deep questions about the coherence of Locke's theory. My goal is not to explore this problem in depth. I would like to point out, however, that it is highly implausible to suppose that Locke did not welcome the view that selves have intermittent existence. I say this, given the fact that his very argument against Descartes appeals to the intermittence of consciousness. A good reading needs to find some way to reconcile this with Locke's pronouncements about two beginnings.[19] What is plain, at any rate, is that the identity between a person at some present moment and a person in the past is secured through consciousness of past actions. And this account works, even in the event that there is an interruption in self-consciousness. Regardless of whether a self continues on through sound sleep, so long as my current self now can

---

[16] *5th Replies*, AT VII.356–357, CSM II.246–247. Citations of Descartes are from Charles Adam and Paul Tannery, eds. *Oeuvres de Descartes* 12 vols. [AT] (Leopold Cerf: Paris, 1897–1910 reprinted, Vrin. 1964–1976), and John Cottingham, Robert Stroothoff, and Dugald Murdoch, eds./transl. *The Philosophical Writings of Descartes*, 2 vols. [CSM] (Cambridge: Cambridge University Press, 1984). Citations made by volume and page.

[17] E II.i.1.9–19, 108–116.

[18] E II.xxvii.1, 328.

[19] One solution involves supposing that Locke's notion of a 'self' is not ontological at all; perhaps it is only an ethical or forensic notion.

remember have doing certain actions before to sleeping, the person I am now is identical to the person who performed those actions earlier.

Yet, despite his efforts to separate the self from substance, Locke nonetheless adopts a model of consciousness that is informed by a specific ontological framework. In order to appreciate this, however, we need to first recognize that Locke adopts the view held by many of his contemporaries, that thought is essentially reflexive. He speaks of "… consciousness, which is inseparable from thinking, and as it seems to me essential to it: It being impossible for any one to perceive, without perceiving, that he does perceive."[20] Indeed, "… thinking consists in being conscious that one thinks."[21] I take it, therefore, that Locke distinguishes between consciousness (in which thinking itself *consists*) and reflection (the second major source of ideas). According to Locke, the latter involves a more robust turning reflection upon oneself; children do not make mental operations their objects of thought until "pretty late" and some people never acquire any clear, perfect ideas at all.[22] By contrast, the former is constitutive of thought itself and is therefore ubiquitous.

This instances a fairly common view at the time which distinguishes between the consciousness which is inseparable from thought, and the second-order examination of one's thoughts through a second act of perception. The distinction is best drawn by Arnauld who says of the former, "… our *thought* or *perception* is essentially reflexive upon itself; as it is expressed more happily in Latin, *est conscia sui*, for I never think without knowing that I think." He calls this reflection "virtual" ("*réflexion virtuelle*") in contrast to a more *explicit* reflection ("*réflexion expresse*"), "… in which we examine our perception by another perception …."[23] This virtual reflexivity of thought is simply consciousness that one is thinking (when one is thinking). And for Locke, thinking itself *consists* in this consciousness. That is: To think is to be conscious that one thinks. I'll use the expression "essential consciousness" to refer this ubiquitous, thought-constituting consciousness.

Now one might also analyze this essential consciousness more deeply by asking: What are the elements of consciousness? What is one conscious of when one is conscious that one is thinking? According to a well-received view, part of this involves consciousness of one's thoughts. A commitment to the reflexivity of thought involves the view that thoughts are essentially elements of consciousness; they cannot exist *except* as elements of consciousness. Descartes explains that "… we cannot have any thought of which we are not aware at the very moment when it is in us."[24] Locke writes: "Our being sensible of it is not necessary to any

[20] E II.xxvii.9, 335.

[21] E II.i.19, 115.

[22] E II.i.8, 107.

[23] Arnauld, *Des Vrayes et des Fausses Ideés*, abbr. (Cologne 1683, p. 46); *On True and False Ideas*, transl. Elmar J. Kremar (Lewiston, ME: The Edwin Mellen Press, 1990), pp. 25–26.

[24] *4th Replies*, AT VII.246, CSM II.171.

thing, but to our thoughts; and to them it is; and to them it will always be necessary, till we can think without being conscious of it."[25] In other words, the view in question is that the very existence of a thought *consists in its being present to consciousness*.

Essential consciousness also involves consciousness of one's own existence, however. In the model of consciousness endorsed by Locke, one is not merely conscious of one's various mental activities; one is also aware of one's very existence: "As for *our own Existence*, we perceive it so plainly, and so certainly, that it neither needs, nor is capable of any proof."[26] For Locke, consciousness that one exists is the same as consciousness that one is a thing.[27] Moreover, such awareness is inherently reflexive. To be conscious that one exists involves an awareness of oneself as a thing. We can say that to be conscious of one's existence is all one with consciousness of *oneself* (*qua* existing thing).

Locke's model of consciousness, however, is structured by a prevailing ontological scheme. The ontological framework is captured by the scholastic distinction between substance and accident. And while other terms can be used to refer to the latter, (such as 'affection', 'modification', 'mode') and some of the ontological details may to some degree differ, the overall picture is clear. It portrays a contrast between a thing and the various states or properties that that thing possesses. The division is to some degree connected to the subject–predicate grammatical division where the latter may be viewed as something like properties, property-instances, or in some way bound up with subject-property ontology. They are "adjectival" on the subject that bears them. And while it is clear that Locke raises serious worries about whether we possess a positive idea of substance, he does not reject this distinction between substance and accident.[28] It is no surprise then that Locke's model of consciousness involves (1) mental acts; (2) oneself *qua* subject of those acts. For Locke, consciousness of one's existence is connected to one's consciousness of mental activities insofar as the latter draws attention to the former: "If I doubt of all other Things, that very doubt makes me perceive my own *Existence*."[29] And:

> If any one pretends to be so sceptical, as to deny his own Existence, (for really to doubt of it, is manifestly impossible,) let him for me enjoy his beloved Happiness of being nothing, until Hunger, or some other Pain convince him of the contrary.[30]

This direction of attention derives from the fact that the variable elements of consciousness, for Locke, are one's own mental states. As such, they inevitably draw attention to one's own existence. Locke endorses, then, a model of essential

[25] E II.i.10, 109.

[26] E IV.ix.3, 618.

[27] E IV.x.2, 619.

28 E II.xiii.17–20, 174–175.

[29] E IV.ix.3, 618.

[30] E IV.x.2, 619–620.

consciousness, informed by substance-mode ontology, which contains two elements: one variable (the mental act) and the other constant (one's own existence) that are connected together as acts to the existence of the thing which has those acts. I call this model of essential consciousness *singular* insofar as both structural elements together constitute *self*-consciousness (that is consciousness of one's states or properties and consciousness of one's existence). As we shall see in what follows, however, Berkeley is led to a transformed conception of essential consciousness which enables him to rehabilitate the notion of a spiritual substance in a way that derives principally from Locke's very notion of the self.

2 Berkeley on the Self and Self-Consciousness

An ontological view can shape one's conception of consciousness, and a shift in ontological framework leads Berkeley to a different model of consciousness. In my view, Berkeley does not view the variable elements of essential consciousness as modes, modifications, properties, states, or the like. Instead, he sees them as things in their own right. While Berkeley's views about substance are not always clear, he is explicit that sensible things exist in the mind not by way of mode, but as a thing perceived in that which perceives it.[31] Moreover, he is fairly hostile to the traditional relation of inherence which is supposed to obtain between substance and mode/accident.[32] We can situate the emergence of these "non-adjectival" mental things within a transformation in conception of consciousness.[33]

More specifically, this transformation concerns various items such as "acts of sensation." A sensation may be viewed as a mental state and as an element of consciousness. If it is no longer viewed as a mental state, however, there is no reason why it cannot still be viewed as an element of consciousness. Locke lists hunger and pain among the thoughts that he is conscious of when he is conscious of himself.[34] But it is not so clear why this hunger and pain ought to be viewed as acts or states of mind, rather than simply objects of which one is conscious. In the interpretation I advocate, both Berkeley and Hume reject the view that such items are mere accidents, modes, modifications, states, and the like which are "adjectival" on a subject, while retaining the view that they are variable elements of consciousness.

In order to better understand what I have in mind, consider the mental objects that one can produce through the use of imagination (such as the image of

[31] PHK I.49; 3D III.237.
[32] PHK I.16–17; PHK I.49.
[33] For an argument that Berkeley rejects mode ontology, see Bettcher *Berkeley's Philosophy of Spirit: Consciousness, Ontology and the Elusive Subject* (London: Continuum Press, 2007), pp. 26–40.
[34] E IV.ix.3, 618.

a tree, an imaginary smell or sound, and so forth). These "items" which one produces needn't be viewed as "adjectival" on the subject. Rather, they can be viewed as things in their own right—things to which a perceiving subject is related through conscious experience.

The view can be further illuminated by turning to two 20th century treatments of Hume's views on self-consciousness. In "On the Observability of the Self" Roderick Chisholm argues that one of David Hume's mistakes is to commit to the existence of ontologically peculiar items identified (by Chisholm) as "sense data" or "appearances."[35] These items are mental "things," distinguished, in part, by their *not* being mental modifications, affections, or properties. By taking these to be objects of "introspective awareness," Hume is led to deny that any subject can be perceived. If he had (correctly, in Chisholm's view) taken appearances as modifications, then he could not have denied that the self can be perceived—*for in perceiving one's modifications and affections, one thereby perceives oneself*. It is precisely the fact that these appearances are not viewed as "adjectival", thinks Chisholm that leads to the (alleged) view that the self cannot be perceived.

Similarly, in "Introspection and the Self" Sydney Shoemaker suggests that lurking behind Hume's denial that anything beyond perceptions can be perceived, is "the act-object" conception of mental states.[36] Accordingly, mental states are mere *perceptual relations* between a subject and mental particulars such as pains and after-images. If all mental states turn out to be relational in this way, however, one will not be aware of any *intrinsic* properties or states of oneself. To be sure, Shoemaker recognizes that this act-object conception of mental states is not one that is ultimately endorsed by Hume (since it presupposes a relation between subject and mental object). However, he thinks that it is presupposed by Hume's concerns about self-consciousness. Berkeley, however, seems like a good candidate for the model that Shoemaker has in mind. But, Shoemaker argues, we really ought to view pain and the like as "adjectival on" a subject of experience. In this way, *to be aware of one's experiences is to be aware of oneself*.

Because Chisholm and Shoemaker write in partial reaction to the sense-data theories of the early 20th century, they are distracted by showing that conception is incorrect. Consequently, both underplay the fact that there is actually a departing from an older view of consciousness (quite similar to that of Shoemaker and Chisholm). Locke, for example, when speaking of consciousness of oneself, speaks of mental acts (of sensation, reasoning, thinking).[37] Rather than viewing the position that Hume and Berkeley hold as simply erroneous, it is perhaps more

---

[35] Chisholm (1969), reprinted in *Self-Knowledge*, ed. Quassim Cassam (Oxford: Oxford University Press, 1994), pp. 94–108.

[36] Sydney Shoemaker, "Introspection and the Self" in *Studies in the Philosophy of Mind*, eds. Peter A. French, Theodoure E. Vehlin, and Howard K. Wettstein (Minneapolis: Midwest Studies in Philosophy 10, 1986), pp. 101–120, reprinted in Cassam (1994, pp. 118–139).

[37] E IV.ix.3, 619.

instructive to see it as an important development in the history of early modern philosophy.

My view, then, is that in Berkeley's model of essential consciousness, while there are two elements of consciousness (both constant and variable), the latter is not a mere modification or state. It is an object in its own right, not adjectival on the perceiving spirit unlike the mental acts of Locke's model.[38] One possible concern with my view, however, is that this shift seems so subtle that it is of little consequence. Why not simply *call* the variable element a mental state (modification, accident, or the like)? After all, it is still dependent on the mind for its existence.

In response, I observe that sheer ontological dependence is not sufficient to yield the view that the dependent items are modes, accidents, property-instances, states, or the like. The simple fact that human minds are dependent upon God for their existence surely does not lead necessarily to Spinoza's notorious view that minds are in some way *modifications* of God. They can still be things in their own right and they need not be "adjectival" on God. Surely a dependent item is not *necessarily* "adjectival" on the thing that supports it. And once these mental items are viewed as things in their own right, they themselves are no longer viewed as mental states.[39] Instead, a subject's *being suitably related* to the item is now more properly viewed as the mental state. This is to say: The thought one produces is not *itself* the mental state. Rather *having* the thought is the mental state.

While this distinction is admittedly subtle it has at least two highly significant interpretive consequences. And in fact, I would argue that these two consequences are *so* surprisingly significant that my interpretation can be defended by appeal to its explanatory power in light of these consequences (coupled, of course, with its close agreement with the text). For the rest of this section, then, I will draw out these interpretive consequences, thereby rounding out my account of Berkeley as well as defending the reading.

The first interpretive consequence is that Berkeley's notorious views about knowledge of spirit can now be fairly easily explained. In the model endorsed by Locke (and advocated by 20th century philosophers such as Chisholm and Shoemaker), consciousness of these variable elements is consciousness of oneself insofar as they are "adjectival" on the subject. As Chisholm suggests, to perceive these states is *ipso facto* to perceive oneself. As I mentioned earlier, in this view essential consciousness is *singular* insofar as it is entirely a form of

[38] I really have two distinct changes in mind. First, for Berkeley ideas are things in their own right. By this, I mean that they do not have 'thing' extended to them with qualification (unlike accidents, for example). On this point, see 3D III.234. Second, the things are non-adjectival on spirit. This is just to say that they are not properties or property-instances which are predicated of spirit. In the interest of brevity, I have avoided a more complex examination of Berkeley's ontological departure.

[39] I suppose one might continue to call non-adjectival, dependent things "modifications." But now it seems that the dispute is about a word alone. What is important for my purposes is that the Berkeleian items which depend upon spirits for their existence are (i) non-adjectival; and (ii) things in their own right.

self-consciousness (i.e. a consciousness of one's mental properties and one's existence).

In Berkeley's model, however, variable elements of consciousness are not "adjectival" on the subject. Instead, they are treated as things in their own right. As a consequence, essential consciousness is now *bifurcated* into two modes of existence: Self-consciousness (of one's own existence) and consciousness of things that are distinct from oneself (one's ideas). Because variable elements of consciousness (ideas) are not "adjectival" on the subject, it is no longer true that to *perceive them is to thereby perceive oneself*. They are no longer to oneself what grass-bending is to grass, and they no longer draw attention to one's existence in the same way. In an effort to keep separate consciousness of ideas from awareness of self, therefore, Berkeley denies that spirits can be *perceived*.[40] Consequently, he rejects Locke's view that one *perceives* one's existence.

Once one takes these mental things as *distinct* from oneself by refusing to view them as mental states or properties, it also becomes clear why they don't have much of a role to play in augmenting one's own self-knowledge (except relationally): One is aware that one exists and that one is *distinct* from these items.[41]

To be sure, regardless of how one analyses consciousness that one smells an odour, it is still true that this consciousness is *self*-consciousness. However, in the Lockean view this consciousness is constituted by consciousness of one's own existence and consciousness of a mental act or state (smelling an odour). By contrast, in Berkeley's bifurcated model it is constituted by consciousness of one's own existence and consciousness of the thing itself (the odour). For Berkeley the odour *itself* is the variable element of consciousness, and "smelling" is nothing but the relation between the two elements of consciousness (the "I" and the odour). In Locke's model, awareness of the variable element is itself a form of *self*-awareness, since it is nothing but a *state of self*, while in Berkeley's model awareness of the variable element is not a form of self-awareness, since the element is something other than oneself—namely, the very odour that one smells.

Moreover, while to be aware that one first smells an odour, and then hears a sound, and then sees a colour must be viewed as *self*-awareness in the two competing accounts, in the Lockean model one is conscious of three distinct mental states (the smelling, the hearing, and the seeing) as variable elements of consciousness. However, in Berkeley's view, the variable elements of consciousness are not distinct mental states at all. The mental states are nothing but the *relations* of consciousness which obtains between the "I" and its objects. Changes in mental state are therefore not intrinsic changes, they are only relational. And while consciousness of one's perceptual history may count as a form of self-consciousness, it does not deliver any deep knowledge of the intrinsic states of the mind.

---

[40] PHK I.27; 3D III.232.

[41] 3D III.234.

Because of this, Berkeley denies that we can have an idea of spirit which represents it by way of resemblance.[42] For suppose that variable elements of consciousness (ideas) are "adjectival" on the self. Then, given that ideas can resemble other ideas, an idea can at least resemble a spirit partially by resembling one of its acts, states, or properties. In such a view, ideas *can* provide significant information about spirits by providing information about the intrinsic states of spirit (namely, the ideas themselves). In denying that ideas can resemble spirits, Berkeley means (in part) to insist that ideas cannot be viewed as modifications of mind. And this is why he may also have a response to Malebranche's claim that one lacks an idea of the mind which can provide information in advance about all of the modifications of which the mind is capable.[43] If, as in Berkeley's view, variable elements of consciousness are not viewed as modifications of the mind in the first place, then the hope for such information is simply based upon an initial conflation of two distinct modes of awareness.[44]

The second significant advantage of this interpretation is that it can now be explained how Berkeley retains spiritual substance within his system. For it has been traditionally worried how Berkeley can defend spiritual substance while arguing that material substance must be rejected. Such parity of reasoning concerns, articulated by Hylas in the 1734 edition of the *Three Dialogues*, have often been viewed as serious and probably fatal to Berkeley's account of substance.[45] Indeed some commentators have supposed that Berkeley may not believe in spiritual substance, or that his appeal is half-hearted.[46] In the view I

---

[42] More accurately, Berkeley denies that ideas can represent spirits because while the latter are active the former are merely passive and inert (PHK I.27). However, he also considers the possibility that spirits possess some *other feature* which enables resemblance and then denies there is such a feature (PHK I.137–138). Thus, Berkeley's denial that ideas resemble spirits is dependent on his claim that they have "nothing in common" (PHK I.89). For a more detailed discussion of these issues see Bettcher (2007, pp. 39–43), and Bettcher "Berkeley on Self-Consciousness" in *New Interpretations of Berkeley's Thought*, ed. Stephen H. Daniel (Amherst, MA: Humanity Books, 2008), pp. 179–202, esp. 184–188.

[43] See *Elucidations of the Search after Truth* XI, pp. 634–635. Citations are from *The Elucidations of the Search after Truth*, transl./ed. Thomas M. Lennon (Cambridge: Cambridge University Press, 1997).

[44] For a more complete account of this idea, see Bettcher (2007, pp. 48–50).

[45] For a defense of Berkeley's conception of spirit, see Bettcher (2007).

[46] Colin M. Turbayne (1959), "Berkeley's Two Concepts of Mind," *Philosophy and Phenomenological Research* 20 (1):85–92; and (1962), "Berkeley's Two Concepts of Mind Part II," *Philosophy and Phenomenology Research* 22 (3): 383–386. Robert Muehlmann, *Berkeley's Ontology* (Indianapolis, IN: Hackett, 1992), pp. 170–204 and his "The Substance of Berkeley's Philosophy" in *Berkeley's Metaphysics: Structural, Interpretive, and Critical Essays*, ed. Muehlmann (University Park, PA: Pennsylvania State University Press, 1995), pp. 89–105. According to Muehlmann, Berkeley conceals a proto-Humean conception of mind by a published appeal to spiritual substance serving none of its traditional functions. Stephen Daniel argues that spirit is not a thing but the sheer existence of its ideas. See, for example, his (2000) "Berkeley, Suarez, and the 'Esse-Existere' Distinction," *American Catholic Philosophical Quarterly* 74 (4): 621–636. For a critique of Muehlmann and Daniel see Hight and Walter (2004), "The New Berkeley," *Canadian Journal of Philosophy* 34: 1–24.

propose, however, while Berkeley rejects mode ontology completely (and therefore does not view spirits as substances in the traditional sense of the basic thing in which modes, modifications, properties and the like inhere), he nonetheless endorses the view that spirits support ideas merely be perceiving them.

Happily, this reading answers another long-standing concern that Berkeley does not have any way to elucidate the mind-dependence of sensible things.[47] The worry is that Berkeley explicitly denies that sensible things exist in spirit by way of mode, property, or attribute.[48] Instead, they exist in spirit as a thing perceived in that which perceives it.[49] The problem is that perception does not itself seem to constitute a relation of dependence. Once we understand Berkeleian perception as the essential consciousness of thought, however, we can see why he might nonetheless think that objects of such consciousness are mind-dependent.[50] In this view, thoughts cannot exist except as elements of consciousness. Their very existence consists in being perceived. In claiming that sensible things are mind-dependent, Berkeley means that they are variable elements of consciousness. They cannot exist except as object of consciousness. For Berkeley, essential consciousness involves two elements: The "I" and the non-adjectival object. This means that sensible things are ever bound to an "I". This is to say: They are mind-dependent.

Because of the fact that Berkeley is now in a position to make sense of support in this way, he is also in a position to transform Locke's notion of the self into a new conception of substance. Notably, Berkeley generally does not use the term 'self,' although he introduces spirit: "This perceiving, active being is what I call *mind, spirit, soul* or *my self*."[51] He writes: "What I am my self, that which I denote by the term I, is the same with what is meant by *soul* or *spiritual substance*,"[52] suggesting that self (and self-consciousness) provide the notion of a *spiritual substance* with content. That Berkeley takes such a strategy seriously emerges in the earliest entries of his notebooks: "Nothing properly but persons i.e. conscious things do exist, all other things are not such much existences as manners of ye existence of persons[;]"[53] and "Qu: about the Soul or rather person whether it be not completely known."[54] Berkeley later eschews use of the term 'person' in

[47] See Raynor (1990, p. 235).

[48] PHK I.49; 3D III.237.

[49] *Ibid*.

[50] Atherton (1983) also draws on this conception of consciousness to illuminate Berkeley's views about spirit. However she unnecessarily concludes from this that ideas, for Berkeley, are "ways of perceiving, not discriminable objects." See her "The Coherence of Berkeley's Theory of Mind," *Philosophy and Phenomenological Research* 43: 389–399, esp. pp. 394–395.

[51] PHK I.2.

[52] PHK I.139.

[53] PC 24. Given Berkeley's explanation, I think he uses 'person' in the way in which Locke uses 'self.'

[54] PC 25. The significance of Berkeley's + sign is controversial. A.A. Luce suggested the mark was used as an obelus to black-list discarded entries. This has been disputed by Bertil Belfrage. See Luce (1970), "Another look at Berkeley's Notebooks," *Hermathena* 110: 5–23 and Belfrage,

part to avoid offending "Church-men."[55] But this doesn't undermine the importance of *self* in grounding Berkeley's notion of spiritual substance.

Now to the degree to which Berkeleian spirit can be counted as a genuine substance, it may seem related to the Cartesian mind, which is bound up with consciousness. For Descartes, the term 'thought' is to "… include everything that is within us in such a way that we are immediately aware of it[;]"[56] a mind is that in which thought immediately resides; and thought itself constitutes the essence of the mind. The view that Berkeleian spirit is something like a Cartesian mind is tempting.[57] Yet it overlooks some of the peculiarities of Berkeley's account of substance as well as the decidedly *Lockean* considerations at work in Berkeley's early notebook formulations.

However, while Berkeley indicates his use of the Lockean notion of *person* in his notebooks, it is also clear that Berkeley rejects the central piece of Locke's theory, namely the determination of identity over time through consciousness of past actions. This can be seen as early as PC 200:

> + Qu: wherein consists identity of Person? not in actual consciousness, for then I'm not the same person I was this day twelvemonth, but while I think on wt I then did. Not in potential for then all persons may be the same for ought we know.

And later, in *Alciphron* (1732), Berkeley has Euphranor, provide the classic critique (later famously provided by Reid) that Locke's theory violates the transitivity of identity.[58] In general, Berkeley rarely discusses consciousness of past actions or even memory, and he never provides any sustained analysis of such mental operations.[59]

That being said, the problem of intermittent existence must remain an issue of concern. At PC 83 Berkeley writes: "Men die or are in state of annihilation oft in a day." It is plausible to read this as a remark about the fleetingness of momentary consciousness or the interruptions that occur during periods of unconsciousness. If this is correct, then Berkeley's early views appear to start from Lockean insights, while jettisoning Lockean consciousness of past actions. Instead, Berkeley admits a fleeting self which is grounded only in momentary consciousness of oneself.

"A New Approach to Berkeley's *Philosophical Notebooks*" in *Essays on The Philosophy of George Berkeley*, ed. Ernest Sosa (Dordecht, the Netherlands: D. Reidel, 1987), pp. 217–230.

[55] PC 713–714. For a classic defense of the view that the Berkeleian spirit is a Lockean person, see I.C. Tipton, "Berkeley's View of Spirit" in *New Studies in Berkeley's Philosophy*, ed. Warren E Steinkraus (New York: Holt, Rinehard, and Winston, 1966), pp. 59–71.

[56] *2nd Replies*, AT VII.160, CSM II.113.

[57] For a defense of this view, see Beardsley (2001), "Berkeley on Spirit and Its Unity," *History of Philosophy Quarterly* 18 (3): 259–277.

[58] ALC VII.viii.299. The argument is apparently first formulated by Henry Grove. See Raymond Martin and John Barresi, *The Naturalization of the Soul: Self and Personal Identity in the Eighteenth Century* (London: Routledge, 2002), p. 72.

[59] For an attempt as such an analysis (although brief), see Bettcher (2007, pp. 83–87).

Yet Berkeley's views on the issue of identity over time develop significantly. Later, Berkeley appeals to the will as some sort of solution.[60] "It seems that the Soul taken for the Will is immortal, Incorruptible."[61] The nature of the identification between soul and will and the role of the will in solving these problems are complex; I avoid any deep investigation here.[62] The important fact for our purposes is that time is pivotal in Berkeley's response to the problem of intermittent existence. At PHK I 98 Berkeley argues that time is "nothing abstracted from a succession of ideas" and that the duration of a finite spirit is to be measured by this succession of ideas in that particular spirit. He concludes that the soul always thinks because no time will pass for the finite spirit during sound sleep.

Berkeley radicalizes Locke's view on the origin of the idea of duration. According to Locke, we obtain the ideas of succession and duration from "Reflection on the train of *Ideas*, which we find to appear one after another in our own Minds."[63] The notion of duration is derived against the backdrop of succession and Locke assures us that should a man have only one idea, he would have no perception of duration, just as he has no perception of duration through sound sleep.[64] Yet Locke allows that time continues while one is sleeping and that the idea of duration can be applied in such cases by appeal to regular natural occurrences.

To be sure, Berkeley's subjectivist theory of time has been vigorously criticized by commentators.[65] My point is not to defend it here.[66] Rather, I contend that Berkeley's views are shaped by the Lockean view of the self, and in particular, the observation that consciousness ceases during periods of sound sleep. Berkeley can ultimately agree with the Cartesian thesis that the soul always thinks only because he has taken the Lockean account as his starting point, and modified it accordingly by appealing to his account of time.

That said, Berkeley still has a serious problem. In underplaying the importance of memory and in identifying time with the subjective succession of ideas, Berkeley still hasn't explained how he avoids a self that is fleeting with each passing thought. For Berkeley, the term "spirit" is given content by appeal to one's own momentary consciousness of self. Yet in the Lockean view, to think is nothing but to be conscious that one thinks where individual thoughts themselves

---

[60] PC 194a.

[61] PC 814.

[62] In my interpretation, Berkeley's final view is that the soul is the will insofar as it is an active being (i.e., an agent). For an account of the simplicity of this active being in the face of its passivity in sense-perception, see Bettcher (2007, pp. 71–80).

[63] E II.xiv.4, 182.

[64] *Ibid*.

[65] I.C. Tipton, *Berkeley: The Philosophy of Immaterialism* (London: Methuen, 1974), pp. 272–292; George Pitcher, *Berkeley* (London: Routledge & Kegan Paul, 1977), pp. 203–211.

[66] For my defense, see Bettcher (2007, pp. 95–101).

seem to have two elements: the variable one (the mental act) and the constant one (one's existence). Given that individual thoughts are fleeting, why doesn't it make sense to say that the "I" which gives "spirit" content is itself fleeting with each passing thought? How can Berkeley actually deliver an "I" that is unified over time rather than a mere sequence of "I"s? The answer, I argue, is once again provided by Berkeley's altered model of essential consciousness. According to Berkeley, 'spirit' can be given content by appeal to one's awareness of oneself—awareness which is distinct from the perceptual awareness that one has of ideas. This is important because it means that while the ideas one perceives are fleeting, the awareness of oneself is not. In the older model of consciousness, to think is just to be conscious that one is thinking. This consciousness involves both one's mental modifications and one's existence. Since consciousness of the latter is a constituent of thought, there will be a new representation with each passing thought, and consequently momentary consciousness will yield a fleeting self. However, in Berkeley's view, thinking is a relation between oneself and one's various objects. While one's objects are fleeting, there is no similar reason to suppose that one's existence is fleeting. Rather, it is a constant against the backdrop of non-intrinsic, relational changes effected by the fleeting ideas. In part because of this, spirits and ideas have very different relations to time. Spirit has duration through the successive ideas it perceives. By contrast, ideas as fleeting constitute the very succession that makes the duration of spirit in time possible. As a result there is a kind of self-consciousness which survives change and continues on indefinitely. Thus Berkeley can allow that spiritual substance remains self-identical over time.

3 Hume on the Self and Self-Consciousness

Berkeley is led to deny that spirits can be perceived. They cannot be presented by ideas (through resemblance). And, as a consequence, there is an important way in which spirits are not objects of understanding at all. One way to put the point is this: While spirits can perceive objects (and in so doing be aware of themselves *qua* existent perceivers), they themselves can never *be* such objects of perception. The gap between spirit and idea for Berkeley is therefore very wide. He claims that they have nothing in common at all but the name—even their respective relations to time are different. He writes: "It is therefore necessary, in order to prevent equivocation and confounding natures perfectly disagreeing and unlike, that we distinguish between *spirit* and *idea*."[67] Further, "[t]here is nothing alike or common in them: and to expect that by any multiplication or enlargement of our faculties, we may be enabled to know a spirit as we do a triangle, seems as absurd

[67] PHK I.139.

as if we should hope to *see a sound*."[68] And given Berkeley's account of natural science as the study of regularities in the phenomena,[69] it is plain that spirits cannot be objects of scientific investigation. Instead, their role is restricted to metaphysics (as first philosophy).[70] It is little wonder, then, that Hume should be interested in rejecting Berkeley's account. It is completely at odds with his own attempt to do for the mind what Newton did for nature.

In what follows, I outline part of what I take to be Hume's strategy for addressing Berkeley's radically bifurcated model of the world (and our knowledge thereof). While Hume, like Berkeley, does not view the objects of consciousness as mental modifications but as mental things, he also holds that these items are constituents of the mind. This means that knowledge of our perceptions is ultimately knowledge of our mind; it means that in perceiving these perceptions, we thereby perceive the mind (or at least a constituent thereof).[71] In this way, Hume hopes to address Berkeley's claim that minds (spirits) cannot be perceived.

Hume's writing suggests the interchangeability of the terms 'self' and 'mind.'[72] Yet while he thinks that both 'self' and 'mind' pick out a bundle of perceptions, this does not mean that Hume assumes interchangeability. Instead, he argues for the thesis that self and mind are one.[73] Hume's notion of *mind* has an ontological import that *self* doesn't. Mind plays the role that spiritual substance might have played (had Hume believed in it). By contrast, an entity is a self only insofar as it is *self*-conscious. It is one thing to say that a mind is a bundle of perceptions, and another to say that the bundle is conscious of itself. For Hume to provide an account of the self, therefore, he must provide an account of how the mind is self-consciousness. In claiming the self is a bundle of perceptions, his point is that to be conscious of oneself is to be conscious of the mental bundle.

Notably, 'self' is scarcely used at all let alone seriously discussed until the rather Lockean entitled section "Of Personal Identity."[74] This section, which proceeds with a discussion of consciousness, contrasts with the previous one which treats of the controversial (but more traditional) question concerning the soul's immateriality. This contrast reflects the way in which Locke's notion of a self replaces more traditional concerns about soul and substance. Moreover, Hume goes on to distinguish his discussion of our propensity to make attributions of

---

[68] PHK I.142.

[69] PHK I.108.

[70] DM 71, 275.

[71] One of the mistakes is to treat perceptions as if they continued to exist independently of our perception of them. Here the issue is not whether the items are capable of existing mind-independently, but whether they do so exist.

[72] This assumption of interchangeably is generally taken for granted.

[73] T Appendix.xvi.399; SBN 635. Penelhum (1955) distinguishes Hume's use of 'mind' and 'person' on different grounds. See "Hume on Personal Identity," *Philosophical Review* 64: 571–589.

[74] Donald Ainslie correctly notes that many commentators have erroneously supposed this section to concern self as substance ("Hume's Anti-*cogito*" unpublished ms.).

personal identity into "our thought or imagination" and "our passions or the concern we take in ourselves."[75] Again, this takes up the related Lockean notions of consciousness (on the one hand) and concern (on the other).

Hume refers to the mind as soul. In the Abstract, the mind is referred to as "the soul, as far as we can conceive it …[,]"[76] and after Hume explains the "true idea" of the mind as a system of perceptions, he proposes that the *soul* is best compared to a republic.[77] Moreover, Hume does not appeal to the argument he uses to show substances (in general) and minds (in particular) are bundles in order to show the self is a bundle. That particular argument involves his principle that ideas be traced to impressions, and the empirical claim that the impression of the target item cannot be found.[78] While Hume uses this type of argument to show that we lack an idea of a self as simple and continued, he does not take this as sufficient to conclude that the self is a bundle of perceptions. Instead, he offers two additional arguments that we will examine in due course. Thus, it seems that Hume accepts (at least provisionally) the Lockean view that self and mind (i.e. soul) can come apart.

With this in mind, it is perhaps tempting to understand Hume's account of this consciousness strictly in terms of second order reflections (i.e., express reflection). Hume appears to identify consciousness itself with "a reflected thought or perception."[79] Moreover, he identifies his process of "entering intimately into what he calls himself" with "turning reflection upon himself" (which again suggests second order cognition).[80] And while it is not clear how we should understand this "turning reflection," it is plausible to understand it in terms of Hume's later talk of "… reflecting on the train of past perceptions, that compose a mind …."[81] When Hume turns reflection upon himself, is he not "reflecting on the train of past perceptions?" If so, it means that in order to think about oneself, one uses memory and imagination in order to do so.[82,83]

---

[75] T I.iv.6.5, 165, SBN 253.

[76] T Abstract.xxviii.414, SBN 657.

[77] T I.iv.6.19, 170, SBN 261.

[78] T I.i.6.1, 16, SBN 15–16.

[79] T Appendix.xx.400, SBN 635.

[80] Compare T I.iv.6.3, 165, SBN 252 with T Abstract.xv.399, SBN 634.

[81] T Abstract.x.v.400, SBM 635.

[82] Garrett (1981) argues that Hume's phrase "reflected thought or perception" is a reference to memory. See "Hume's Self-Doubts about Personal Identity," *The Philosophical Review* 90 (3): 337–358.

[83] According to Ainslie (2001), when Hume turns reflection on himself, he is interested only in an activity performed by philosophers involving "secondary ideas." (T I.i.11.10, 2, SBN 6) A "secondary idea" (according to Ainslie) is supposed to take a perception itself as an object, while primary ideas share an object in common with an impression. See "Hume's Reflection on the Identity and Simplicity of the Mind," *Philosophy and Phenomenological Research* 63 (3): 557–578. One difficulty with this reading is that it deliberately leaves no room for the common-folk to think about themselves in this way. By contrast, it seems more attractive to construe reflection

As I shall argue there is something right about the view that Humean self-consciousness is a kind of second-order reflection upon one's past self. However, I wish to put such considerations to the side for now. Instead, I want to show how Hume nonetheless explicitly aims to address the virtual reflexivity inherent in essential consciousness. And this requires that he account for both consciousness of perceptions as well as consciousness of one's present existence.

Begin by noticing that Hume also explicitly recognizes a cognitive relation between mind and perception which is unmediated by any second order thought. Accordingly, Hume allows that perceptions (and only perceptions) are "present to the mind."[84] That "presence to the mind" is a cognitive relation is shown by Hume's identification of this relation with "appearance in the mind[:]"

> … [T]he *appearance* of a perception in the mind and its *existence* seem at first sight entirely the same[;] it may be doubted, whether we can ever assent to so palpable a contradiction, and suppose a perception to exist without being present to the mind.[85]

Moreover, to "appear in the mind" is to "be perceived by the mind." Thus Hume speaks of "… our pains and pleasures, our passions and affections, which we never suppose to have any existence beyond our perception …."[86]

The fact that this cognitive relation is not mediated by a second-order perception is evidenced by Hume's account of it. In order to show how it is possible that the appearance in the mind and the existence of a perception can come apart, he claims that the mind is nothing but a heap of causally connected perceptions. So a perception can be separated off from the bundle which it helps to constitute.

He then goes on to explain how we "… conceive an object to become present to the mind, without some new creation of a perception or image; and what we mean by this *seeing*, and *feeling*, and *perceiving*."[87] This indicates that Hume aims to account for how objects become present to the mind *without the mediation of an additional perception*. However, Hume has already warned us that he is using 'object' and 'perception' interchangeably.[88] In effect, Hume is providing an account of how perceptions become "immediately present to the mind" in a way that does not involve another perception. His account is that "external objects" become present to the mind (are seen, felt, etc.) insofar as they have a causal impact upon the bundle (by augmenting present reflexions and

---

as involving only memory and imagination (something to which the common-folk ought to be able to appeal). Notably, while memory plays a central role in Hume's account of personal identity, secondary ideas are not mentioned even once. For further critique of Ainslie's view, see A.E. Pitson, *Hume's Philosophy of the Self* (London: Routledge, 2002), pp. 75–80.

[84] T I.ii.6.8, 49; SBN 67.

[85] T I.iv.2.37, 137, SBN 206.

[86] T I.iv.2.16, 129, SBN 194.

[87] T I.iv.2.38, 137, SBN 207.

[88] T I.iv.2.31, 134, SBN 202.

passions, storing memory with ideas[89]). "Appearance in the mind" is not understood in terms of the production of a specific second-order perception which takes an original perception as its object; it is understood in terms of the causal impact made by the perception. That is, to become present to the mind is to make a causal impact upon the mind. And Hume himself is prepared to call this a kind of consciousness. He writes:

> For since all actions and sensations or the mind are known to us by consciousness, they must necessarily appear in every particular what they are, and be what they appear. Every thing that enters the mind, being in *reality* a perception, 'tis impossible any thing shou'd to *feeling* appear different. This were to suppose, that even where we are most intimately conscious, we might be mistaken.[90]

He refers to "appearing in the mind," and it is hard not to read "feeling" in terms of the causal impact account. Importantly, there is a clear link to "intimate consciousness."

So it seems that Hume's account of "appearance in the mind" is designed to explain the essential consciousness of thought (virtual reflection).[91,92] It would allow us to construe the "palpable contradiction" that Hume tries to undermine as the well-accepted "absurdity" that thought should exist without consciousness of it. And given that Hume is interested in this section to undermine the Berkeleian view that *esse* is *percipi*,[93] it seems he is reading Berkeley in a particular way. He understands Berkeleian mind-dependence in terms of the "palpable contradiction" that an object of this consciousness should ever exist without being an object of consciousness: There is no such thing as an unconscious pain; there is no such thing as an unconscious thought. And he is explicitly trying to show how such objects can be separated from the consciousness which binds them to the mind. Effectively, then, Hume is taking direct aim at Berkeley's attempt to elucidate the relation of support between spiritual substances and their ideas in terms of essential consciousness. In this way, the common view that because Berkeley is

---

[89] T I.iv.2.40, 138, SBN 207.

[90] T I.iv.2.7, 127, SBN 190.

[91] For interpretations of Humean consciousness, see Thiel (1994), "Hume's Notions of Consciousness and Reflection in Context," *British Journal for the History of Philosophy* 2 (2): 75–115; Waxman (1994), and Stevenson (1998), "Humean Self-Consciousness Explained," *Hume Studies* 24 (1): 95–129.

[92] One might object to this account of consciousness on the grounds that Hume considers a mind which has only one perception. Shouldn't consciousness accompany the perception even then? In my view, Hume is interested to show that the existence of a perception and consciousness thereof can come apart. Since the oyster's perception makes no causal impact, it is effectively an 'unconscious perception.' For an expression of this concern, see Green (1999) "The Idea of a Momentary Self and Hume's Theory of Personal Identity," *British Journal for the History of Philosophy* 7 (1): 103–122.

[93] Robert Fogelin reads the section as an argument against Berkeley. See his *Hume's Skepticism in the Treatise of Human Nature* (London: Routledge & Kegan Paul, 1985).

notoriously incapable of accounting for mind-dependence, his position is simply dismissed by Hume is effectively undermined.[94]

While Berkeley uses the mind-dependence of ideas to show that what the vulgar take to be real things cannot exist mind-independently[95] Hume takes the coherence of the vulgar claim that the items immediately present exist mind-independently to show that perceptions themselves can exist without being perceived.[96] In this way, Hume's intervention leads him to take a position that might seem equally extreme as Berkeley's thesis that external objects are nothing but sensations (i.e. variable elements of consciousness) and are therefore mind-dependent. Because Berkeley sides with the philosophers that ideas cannot exist without being perceived, he believes that vulgar objects of perception cannot exist unperceived. Because Hume sides with the vulgar that it is coherent to suppose that objects of perception exist independently of being perceived, Hume is led to reject the philosophical truism that objects of consciousness (such as pains) cannot exist without consciousness thereof.

This conflict brings an interesting issue to the surface. For Berkeley, the content of 'spirit' is provided by consciousness of one one's own existence (the "I"). His point is that ideas cannot ever exist except as elements of consciousness, and since essential consciousness requires an "I", these ideas cannot exist without an accompanying, perceiving "I". The relationship between the mind and self-consciousness for Hume, however, is a bit tricky. In "Of Scepticism with regard to the Senses," Hume seems to be showing how perceptions can exist independently of the mind itself. Here, there is no discussion of self-consciousness at all. But surely this is an issue Hume needs to address. What is it to be conscious of the *mind*, and what is it to be conscious of *one's own existence*? Hume, it seems, offers a provisional answer in the section "Of Personal Identity."

When Hume famously begins that section, "There are some philosophers, who imagine we are every moment intimately conscious of what we call our Self[,]"[97] he intends to reject consciousness of one's own *existence* as a kind of datum simple and continued over and above one's consciousness of one's various different perceptions. This rejection is of a piece with his departure from the Lockean model of consciousness. Because, like Berkeley, he does not view variable elements of consciousness (perceptions) as modifications, affections, or the like, he does not view them as "adjectival" on some state-bearing mental

[94] According to Raynor (1990, p. 235), the relationship between Berkeleian spirit and idea is fundamentally unclear and Hume saw this as an embarrassment to be solved by the bundled mind. Yet while it has been unclear to many commentators what Berkeley means when he says that spirits *perceive* ideas, may be clarified in terms of consciousness. According to the common philosophical view, no thought can exist without being an object of consciousness awareness (i.e., without appearing to a mind). Hume seems to have fully grasped Berkeley's point, and attempted to dissolve "the palpable contradiction" through his own bundled account of the mind.

[95] 3D III.262.

[96] T Appendix.xiii.399, SBN 634.

[97] T I.iv.6.1, 164, SBN 251.

subject. In light of the fact that this model of essential consciousness has been rejected, it is no longer true to say that there is an additional datum such that in perceiving one's perceptions, one thereby perceives one's existence.

Consider: Is Hume, in reaction to Locke, so skeptical as to deny his very own existence? On the face of it, it is hard to believe. So let's say that he accepts his own existence as self-evident. Then it is clear that somewhere in the *Treatise* Hume would need to provide an account of how it is that we are conscious of our existence. And the closest that Hume ever gets to addressing this issue is in "Of Personal Identity." He must, therefore, be providing his account in this section. We ought to read the opening remarks as an attack on a particular account of how we are conscious of our own existence. Once we accept this (and bear in mind Hume's rejection of the older conception of consciousness), my reading of Hume is unavoidable. What Hume is rejecting in the opening paragraph is precisely that model.

In further defense of this view, let me note that Locke is likely one of Hume's targets.[98] Aside from his use of the term 'self,' Hume considers the issue of proof: "To attempt a farther proof of this were to weaken its evidence …."[99] He likewise takes up this issue of directed attention: "The strongest sensation, the most violent passion, say they, instead of distracting us from this view, only fix it the more intensely, and make us consider their influence on *self* either by their pain or pleasure."[100] It is precisely because Hume rejects the Lockean model of consciousness that he can reject Locke's directed attention.[101] But what *Locke* is interested to discuss is only knowledge of one's own *existence*.

Yet Hume is clear that he is interested in more than the feeling of one's existence; "some philosophers" also claim of the self "… that we feel its existence and its continuance in existence; and are certain, beyond the evidence of a demonstration, both of its perfect identity and simplicity."[102] Consequently, Hume is also interested in the knowledge that the self is both perfectly identical over time as well as simple. This suggests that in addition to consciousness of the self's existence, he is interested in consciousness of *what the self is*.

But Locke explicitly leaves open the question whether the self (or at least the substance which is appropriated to it) is simple or complex. The further question whether the self *itself* is simple or complex makes no sense. For any being is a self so long as it is conscious of pleasure or pain (and therefore its own existence).

---

[98] The identities of the philosophers addressed by Hume are not clear. In his ms., Ainslie argues (in favour of the traditional view) that Descartes is a central target. While I think that Descartes adopts the conception of consciousness that is under attack, I also think that the decidedly *Lockean* nature of section 6 warrants caution.

[99] T I.iv.6.1, 164, SBN 251.

[100] *Ibid*.

[101] For a different account of Hume's response to this directed attention, see Donald Ainslie, unpublished ms. "Hume's Anti-*Cogito*."

[102] T I.iv.6.1, 164, SBN 251.

How can simplicity apply when the issue is simply whether a being is conscious or not? The issue, it seems, concerns the *model* of consciousness in play.

According to Locke, the *idea* of existence is a simple one that is suggested to the understanding by every object it perceives.[103] So if one has an idea of *one's own* existence, for Locke that idea is simple. While Locke himself curiously never discusses the idea of his own existence, we can plainly see that the perception that he has of it (i.e., his consciousness of himself *qua* thing) is simple.[104]

If this is correct, Hume's thesis that the self is a bundle of perceptions ought to be understood, in part, as an analysis of consciousness that one exists. Again, it seems that Hume is showing us something more than this (namely, he is showing us *what* the self is). However, while Locke holds that a simple idea of existence is suggested by every object we perceive, according to Hume, the idea of the thing's existence is the very same as the thing itself.[105] It is therefore little surprise that in order to perceive our existence, we must perceive ourselves.

Let me add here that Hume is explicitly trying to account for the relationship between perception and self. He asks, "After what manner, therefore, do they belong to the self; and how are they connected with it?"[106] In my reading, the question that concerns him is the analysis of essential consciousness that one is thinking. What is it to be conscious of oneself (i.e., one's own existence) and how does this relate to consciousness of our perceptions? Recall that these two elements together constitute Lockean self-consciousness. According to Hume, since we are minds, in order to perceive our existence we must perceive the bundle that we are. To be conscious of one's existence is to be conscious of oneself *qua* mind—a consciousness which *ipso facto* includes within it consciousness of perceptions.

We now return to a question I raised awhile ago. Is self-consciousness, for Hume, a kind of second-order express reflection or is it the kind of virtual reflection characteristic of essential consciousness? We have already seen Hume's attempt to accommodate consciousness of perceptions in a way that does not require further, second-order perceptions. And this suggests a promising line for the mind as well. To perceive perceptions is to perceive oneself. And if the unmediated appearance of a perception in the mind is its causal impact upon the bundle, then the unmediated appearance of the mind (i.e., the self) is the sum total of causal impacts by the perceptions that constitute the mind.

One of the arguments Hume provides in favour of his bundle account of the self further suggests such an analysis. Hume notes that were he in sound sleep, he would be *insensible of himself*. He concludes from this that he would be a perfect non-entity in sound sleep. Likewise were his perceptions removed after death (so

---

[103] E II.vii.7, 131.

[104] And there does seem to be a problem concerning the origin of this idea. If it arises as a consequence of the structure of thought itself, is it not innate?

[105] T I.ii.6.4, 48, SBN 66.

[106] T I.iv.6.3, 164–165, SBN 252.

that he could no longer feel, see, love, etc.) he would be a non-entity. The fact that one is insensible of oneself does not show that one is a non-entity, unless one means a self (which requires self-consciousness in order to exist). And if the sleep renders one *insensible* of oneself, then presumably when one is awake one *is* sensible of oneself. What is it to be *sensible* of oneself? The conclusion pressed by Hume: To perceive one's perceptions is to perceive oneself. For to remove one's perceptions would render oneself insensible of oneself and therefore non-existent.

Alas, the other (more famous) argument Hume provides, is far less clear. Here Hume issues his famous Complaint that whenever he "enters most intimately into what he calls himself" he always perceives perceptions and perceives only perceptions. He concludes that the self is a bundle of perceptions. And his argument is that whenever we do this, we perceive only perceptions. Because of Hume's appeal to "entering intimately into oneself" and his Appendix recasting on this in terms of "turning reflection" it seems appropriate to read this as express reflection.

That said there is nonetheless a noteworthy oddity here. For this "turning reflection" seems actually entirely irrelevant to the argument. After all, Hume has also earlier remarked:

> *Let us fix our attention out of ourselves* as much as possible: Let us chace our imagination to the heavens, or to the utmost limits of the universe; we never really advance a step beyond ourselves, nor can conceive any kind of existence, but those perceptions, which have appear'd in that narrow compass.[107]

Since for Hume, the mind is a bundle of perceptions, regardless of whether one realizes it or not, to perceive a perception is to perceive a constituent of one's mind, and is therefore to perceive oneself. So regardless of whether one "turns reflection" or not, one *always* perceives one's mind.

One possible response is to say that while Hume admits that consciousness of any perception is a consciousness of the thing that turns out to be oneself, this doesn't quite amount to *genuine* self-consciousness. *Genuine* self-consciousness requires that one *know* that the object of awareness is oneself. And Hume *denies* that the vulgar recognize all perceptions as constituents of their own minds. According to Berkeley, the vulgar fail to appreciate that their objects of perception are ideas. This is to say, they fail to appreciate that their objects cannot exist except as objects of a conscious mind. Yet since Hume has separated the existence of a perception from its appearance in the mind, however, this can't be what the vulgar fail to appreciate. What they fail to appreciate is that all of their objects are actually constituents of their mind. But what is required for genuine self-consciousness, one might argue, is something more. And this something more is provided by Hume's explicit turning of reflection upon himself.

Yet this reply isn't quite right. It is true that Berkeley and Hume both accept "the doctrine of the philosophers" that we only ever perceive Humean perceptions/Berkeleian ideas.[108] And it also true that for both this is something that is generally

[107] T I.ii.6.8, 49, SBN 68, my emphasis.

[108] T Appendix.399, SBN 634; 3D III.262.

not recognized by "the vulgar." However the distribution of doctrines is actually more complex. For the vulgar *do* recognize that pains, hunger, depression and the like are mental things.[109] But they just *don't* recognize this with respect to sensible things such as colours and sounds, or figure and solidity. Modern philosophers, by contrast, may recognize that colours and sounds are mental things. However, they may continue to err with respect to figure and solidity.[110,111]

What this means is that even the vulgar know that *some* of the things they perceive belong to their own minds. But this is enough to furnish at least an incomplete but unmediated consciousness of self. Indeed, this actually agrees nicely with Hume's own remark: "and in common life 'tis evident these ideas of self and person are never very fix'd or determinate."[112] And it is certainly enough to yield consciousness of one's own existence. Therefore, I continue to insist, Hume's appeal to "turning reflection" is irrelevant to the argument.

More importantly, however, "turning reflection" doesn't even seem equipped to do the job in the first place. All second order perceptions (memories, secondary ideas of ideas) must be viewed as effects of their object. Given that Hume's account of cause and effect requires temporal contiguity, there is never a way in which a secondary representation could constitute consciousness of one's own existence *in the present*.[113] In order for Hume to provide an account of this sort of self-consciousness, he would need the *unmediated* consciousness of oneself—the appearance of all perceptions to the mind, sufficiently unified to consciousness of one's own present existence.

So given that Hume has already tried to account for unmediated consciousness of perceptions, it seems plausible that he would try to do the same for one's existence. Moreover, one of his arguments for the self suggests such an analysis. The other, by contrast, is strangely complex. The appeal to "turning reflection" runs against this reading, yet the appeal is also both irrelevant and ultimately incapable of yielding consciousness of one's present existence. There is good reason, therefore, to see Hume as attempting to accommodate unmediated consciousness of one's present existence.

A further argument in favour of this reading is that it provides a good way of understanding Hume's announced perplexity about his own account in the

---

[109] T I.iv.2.12, 128, SBN 192; PHK I.3.

[110] T I.iv.2.12–13,128; PHK I.10.

[111] On the face of it this view may raise the concern of how thought can be reflexive, if one is not conscious of these mental things *as* mental things. Yet, Malebranche recognizes sensations as objects of "*sentiment intérieur*," while also arguing that it is unclear from this modality of awareness whether the sensations are modifications of mind or extension. Instead, reasoning is required, involving an appeal to the idea of extension: Since sensations do not belong to extension (as per the idea of it), they are modifications of mind. See *Elucidations of the Search after Truth* XI, pp. 634–635. Citations are from *The Elucidations of the Search after Truth*, transl./ed. Thomas M. Lennon (Cambridge: Cambridge University Press, 1997).

[112] T I.iv.2, SBN 189–190.

[113] For a related remark, see (Stevenson 1998, p. 116).

Appendix. Hume claims that he cannot render consistent the following principles: (1) All our distinct perceptions are distinct existences; and (2) We never perceive any real connexion among distinct existences. What makes them 'inconsistent' is Hume's commitment to the third principle: (3) We are conscious of our own present existence.

In affirming the first principle, Hume claims that perceptions can exist independently. This leads him to argue for the claim that to be conscious of one's existence is to perceive the perceptions that constitute the mind. Yet, the appearances of all perceptions do not themselves constitute perception of oneself, since there are no connections among these existences. How can the perception of a plurality constitute the perception of one's singular existence?

The solution, it should seem, would be provided by the understanding which has the task of comparing ideas. Alas, the understanding never perceives any real connexion among distinct existences. So how, then, does one perceive one's own present existence *at all*? The hope for any kind of genuine momentary consciousness of self is effectively immersed in perplexity.

Strangely, the only possibilities Hume takes seriously are (1) that perceptions are adjectival on a subject (i.e. a Lockean account) or (2) that one can perceive a real bond among the perceptions (i.e. what would appear to be required on his own account). Neither work, since Hume has rejected the Lockean account of self-consciousness, and since his own account of the self as a bundle has likewise been thwarted. Yet this perplexity could be alleviated by admitting Berkeleian consciousness of the "I". For if an analysis of essential consciousness that one is thinking cannot be accommodated in terms of perceptions alone (as Hume intends), then it might seem one ought to affirm an awareness of one's own existence that is distinct from one's perception of objects (as Berkeley does). Yet Hume does not do this. But why not? What about Berkeley?

4 Hume Against Berkeley on the Self and Self-Consciousness

It emerges as a pressing question why, while Berkeley appreciates the considerations which lead him to distinguish self and ideas, Hume, apparently, does not. In answer to this, several points are worth noting. Hume's failure to accommodate consciousness of one's own present being does not undermine his account of personal identity. According to Hume, the true idea of the mind arises from memory, and "thought alone finds personal identity, when reflecting on the train of *past* perceptions, that compose a mind …[;]"[114] this again suggests memory. Certainly, memory is a sufficient sort of consciousness of oneself (the train of past perception) to allow for the generation of the passions of pride and shame, and consequently the idea/impression of oneself (*qua* bundle of perceptions).

[114] T Appendix.xx.400, SBN 635, my emphasis.

In this way, Hume's failure to accommodate consciousness of one's own existence *in the present* hardly constitutes a failure of his account overall. Indeed, his account here seems to be entirely based on reflection upon one's past self, which as we have seen, is incapable of yielding consciousness of one's present existence anyway. Clarifying this helps show that what Hume sacrifices in his account is something that could never be accounted for by appeal to any second-order perception. This is something that he simply cannot have.

Is it not odd that Hume should have failed to notice this? And is it not good fortune that the rest of Hume's account of the self is left intact? Indeed, is it not even better fortune that the resulting perplexity should give Hume more evidence in favour of skepticism? Note that Hume does not pronounce the problem "absolutely insuperable." Others may reconcile the contradiction. But how does one reconcile the two principles that Hume finds problematic? The only solution (and surely Hume recognizes this) is to jettison the third principle which makes them problematic, namely the principle that *we are conscious of our own present existence.*

Such a solution would be radical. It would require coming close to Locke's impossible skeptic who denies his own existence. It would involve denying that we are ever really *conscious* of our own present existence. Instead, the only consciousness available would be of our *past existence* (as revealed through memory). Yet the basis for such a view can already be found in Hume's writing. Hume's very account of "appearance in the mind" involves time delay. A perception appears in the mind insofar as it makes a causal impact. But this occurs *after* the perception comes into existence.

Indeed, Hume's views about time render the very notion of momentary consciousness of present existence problematic. While Hume follows Berkeley in rejecting abstract ideas, his views about time are, although closely related to, importantly different from those of Berkeley. Hume allows (as does Berkeley) that particular ideas can become general by representing a plurality of objects. For Hume this general idea of time is derived from any succession of perceptions whatever; any application of the idea to an unchanging idea is fictional.[115] It is by engaging in this fiction of considering an uninterrupted perception against the backdrop of succession that provides us with the notion of identity (over time).[116] Thus, Berkeley's attempt to find a consciousness of one's own unchanging existence through time is going to be undermined. In order to be aware that one exists in time, the idea of time must be applicable. But it can only apply, according to Hume, in the case of succession. Consequently, the very notion of a kind of inherent consciousness of one's present temporal existence which accompanies all thought appears problematic in Hume's theory.[117]

---

[115] T I.ii.5.28–29, 47–48, SBN 65.

[116] T I.iv.2.29–30, 133–134, SBN 201.

[117] Yet this isn't enough to rule out Berkeleian self-consciousness as such. It only yields the view that consciousness of one's existence is not time-bound.

Notably, because Hume, unlike Berkeley, doesn't identify time with the subjective succession of perceptions per mind, he leaves open the possibility that time elapses while we are in sound sleep. Hume gives the clear sense that he allows for the possibility of time elapsing during sound sleep: "A man in sound sleep, or strongly occupy'd with one thought, is insensible of time …."[118] Instead, Hume's view is that we cease to exist during periods of sound sleep: "When my perceptions are remov'd for any time, as by sound sleep; so long am I insensible of *myself*, and may truly be said not to exist."[119]

Hume rules out the rightful application of the idea of time only to anything not involving succession. This leaves open the possibility that the idea can be applied to an imagined situation in which we are sleeping (so long as we imagine perceptions in succession). Given that Hume has done considerable work showing (*pace* Berkeley) that it is intelligible to suppose perceptions can exist unperceived by a mind, he can also say that it is intelligible to imagine several perceptions succeeding each other independent of our own mind. Thus, Hume's attack on Berkeley's model of mind-dependence affords him a way to affirm the conceivability of a mind that exists while sleeping.

Hume thereby moves closer to the Lockean position which allows for at least the intelligibility of the intermittent existence of a self. The difference is that while Locke attempts to secure genuine identity over time through consciousness of past actions, Hume does not. He specifically rejects Locke's view by claiming that memory does not so much constitute as it does discover personal identity.[120] Instead, Hume's "solution" is to deny that there is any strict identity of a mind over time at all.

Yet memory does work in Hume's theory of personal identity that is analogous to that done by Locke's consciousness of past actions. While memory has a role to play in constituting one of the relations which obtains between our successive perceptions and makes us attribute a perfect identity to the mind (resemblance), it plays a more important role in discovering the relations which make us attribute this identity.[121] Indeed, Hume suggests that it is chiefly from memory that our idea of the identity of the mind arises. So memory is implicated in the very *institution* of the connections themselves:

> We only *feel* a connexion or a determination of the thought, to pass from one object to another. It follows, therefore, that the thought alone finds personal identity, when reflecting on the train of past perceptions, that compose a mind, the ideas of them are felt to be connected together, and naturally introduce each other.[122]

---

[118] T.I.ii.3.7, 28, SBN 35.

[119] T I.iv.6.3, 165, SBN 252. For a dissenting view, see Wayne Waxman, *Hume's Theory of Consciousness* (Cambridge: Cambridge University Press, 1994), p. 222, n. 19.

[120] T I.iv.6.20,171, SBN 262.

[121] T I.iv.6.20,171, SBN 262.

[122] T Appendix.xx.400; SBN 635.

Both memory and imagination are involved in this reflection upon the train of past perceptions; this is what yields the feeling of connectedness that is then confounded with identity. The point is that while memory does not produce personal identity through the appropriation of past actions to a present self, it is involved in the very institution of the "connections" by which the perceptions are grouped together.Thus, Hume remarks: "However extraordinary this conclusion may seem, it need not surprise us. Most philosophers seem inclin'd to think, that personal identity *arises* from consciousness; and consciousness is nothing but a reflected thought or perception."[123]

This helps us explain the apparent irrelevance of "turning reflection upon oneself" in Hume's first argument for his analysis of unmediated self-consciousness. If Hume always intended consciousness of one's present existence to be exiled as problematic, then the only kind of self-consciousness which Hume can allow is consciousness of one's past existence. In such a model, this consciousness will inevitably be facilitated by a second order perception. Moreover "turning reflection" (consciousness of one's past existence) ultimately proves central because it is ultimately the only sort of self-consciousness that Hume will accommodate. It is this unproblematic reflection upon one's past existence which leads to the problematic view that consciousness of one's present existence should be tantamount to the perception of all present perceptions.

With this in mind, it is worth considering that Berkeley occupies a peculiar position in Hume's *Treatise*. While Hume attacks Berkeley's theory at its core when he first introduces his account of mind (in "Of Scepticism with regard to the Senses"), Hume's two main sections which actually concern the soul and self ("Of the Immateriality of the Soul" and "Of Personal Identity") don't address Berkeley's at all. In these sections, Hume aims to show that we lack an idea of the soul as substance and then that we lack an idea of the self as distinct from its perceptions.[124]Berkeley however, would obviously dispute neither of these conclusions.[125] In Berkeley's view, while one is conscious of one's own existence, no idea can resemble and thereby represent a spirit. Furthermore, when Hume defends his own account of self-consciousness by issuing his famous Complaint, he claims that only perceptions can be *observed* or *perceived*. But this is again something to which Berkeley would have agreed.

What is even more striking is that, while Berkeley is not touched by the central arguments Hume uses against competing conceptions of mind and self, Berkeley is the *first* philosopher hinted at in Hume's introduction of his account of the mind "Of Scepticism with regard to the Senses." Yet there, Hume offers his account of the mind *with no argument at all.* More remarkable, Berkeley is once again alluded to in Hume's Appendix treatment of his own account. Here Hume

---

[123] *Ibid.*

[124] For the view that Hume addresses Berkeley and Locke in the opening passages of "Of personal identity" see Pappas (1992), "Perception of the Self," *Hume Studies* 18 (2): 275–280.

[125] For a more complete account see Bettcher (2007, pp. 111–113).

does provide an argument. He says that since chimneys and tables can exist without being present to the mind, so too, can perceptions themselves. Alas, his argument simply seems to beg the question against Berkeley.

These facts together indicate that Berkeley may occupy a special place in Hume's account. Rather than being explicitly argued against (at least in the main body of the *Treatise*), Berkeley's vision of the world seems to be a position that Hume simply seeks to supplant with his own account. As I have been arguing, Hume is interested in responding to a Berkeleian account of the world which splits spirits and ideas into two radically different kinds of things and thereby makes it impossible for spirits to be objects of understanding in an important sense. Elsewhere, I have also argued that the dispute between Hume and Berkeley concerns fundamentally different philosophical project and starting-points.[126] It may then be that Hume is simply rejecting Berkeley as an opening move. But if so, he would still need to accommodate Berkeley's position. But how?

It is sometimes felt that Hume "leaves something out" of his famous Complaint. In particular, it has been supposed that Hume was "looking in the wrong place." In my view, this suspicion can be given content once we recognize that what Hume specifically leaves out is Berkeley's conception of self-consciousness. Hume's Complaint is formulated strictly in terms of *perception* and *observation*. Yet the perceiving "I" is precisely that which *cannot* be perceived in Berkeley's view. While Hume *can* reject any Lockean view that one's present existence (over and above one's perceptions) is perceived because he has rejected the ontology which informs Locke's model of essential consciousness, nothing that Hume says in this passage explicitly addresses Berkeley's unperceived perceiver.

It is also sometimes felt that what Hume "leaves out" is precisely what he ultimately needs to hold all of the perceptions together when he seems to despair in the Appendix. Yet, if I am right that the Appendix perplexity actually leaves Hume's entire account unscathed, then the Appendix remarks serve no other purpose than to place consciousness of one's present existence in jeopardy. It seems to be the case, then, that Hume is actually *addressing* Berkeleian self-consciousness in the puzzling Appendix remarks by effectively throwing consciousness of one's present existence "under the speeding bus" of skeptical perplexity. If so, then what Hume "leaves out" of his Complaint in the *Treatise* (namely Berkeley's unperceived perceiver) is ultimately accommodated by a willingness to "plead the part of the skeptic" through placing consciousness of one's own present existence in extreme jeopardy.

In response to the Berkeleian position, Hume may well be endorsing one of the two following strategies. First, rather than accepting the Berkeleian thesis that it is impossible for ideas/perceptions to exist without a mind, he may be willing to leave his account of momentary self-consciousness perplexed (appealing to the privilege of the skeptic). Second, he may actually endorse the only solution to the

[126] Bettcher (2007, pp. 107–116).

tension between his two principles, namely the rejection of the view that we are conscious of our own existence, and the subsumption of all self-consciousness into memory.

The last solution is interesting because it brings into contrast Berkeley and Hume on consciousness, self, and time. While both are led to new models of consciousness through a rejection of the older ontology, the models of consciousness are contrasting. For Berkeley, consciousness of one's present existence has been separated from consciousness of mental items which are no longer viewed as belonging to oneself. Consequently, this *non-perceptual* consciousness of self survives through consciousness of succession, and in this way exhibits a unity over time. For Hume, one is only *perceptually* conscious of one's past existence through memory which itself helps institute the "connections" by which the idea of oneself arises. Consciousness of one's present existence, however, is "left for dead." So strong, it seems, was Hume's commitment to securing a scientifically perceivable mind, at any rate, that in order to address Berkeleian spirit he dared titter and sway at the precipice of the most of outrageous skepticism.[127]

[127] I would like to thank Roberta Morris, Randall Parker, and Kayley Vernallis for their helpful comments on earlier articulations of these ideas. I give special thanks to Susan Forrest for her invaluable insight and support. This article is based upon ideas originally developed in my dissertation. I thank my dissertation supervisor, John Carriero, for his thoughtful guidance.

Making an Object of Yourself: On the Intentionality of the Passions in Hume

Amy M. Schmitter

1 A General Problem with Intentionality

Shortly after opening Book II of the *Treatise*, Hume presents us with a curious claim. He declares that "pride and humility, tho' directly contrary, have yet the same OBJECT. This object is self."[1] This passage is puzzling on several counts, not least of which is deciding what to make out of this idea of the self.[2] For the penultimate section of Book I, concluded a scant 14 pages before, develops a highly skeptical account of personal identity that seems to deny that we have a well-defined sense of self. Hume there takes aim at those "who imagine we are every moment intimately conscious of what we call our SELF, … and are certain, … both of its perfect identity and simplicity." Instead, he maintains:

> when I enter most intimately into what I call myself, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch myself at any time without a perception, and never can observe any thing but the perception.[3]

True, the skepticism of Book I is directed particularly to "personal identity, as it regards our thought or imagination," rather than "as it regards our passions or the concern we take in ourselves."[4] And in Book II, Hume goes on to qualify this

[1] T II.i.2, 277. References to primary sources are as follows: T = Hume, *Treatise of Human Nature*, 2nd edition, ed. Selby-Bigge (Oxford: Oxford University Press, 1978), cited by book, part, section and page; E = Locke, *Essay Concerning Human Understanding*, (Amherst, NY: Prometheus Press, 1995), cited by book, chapter, paragraph and page; PHK = Berkeley, *A Treatise Concerning the Principles of Human Knowledge* (Indianapolis, IN: Hackett, 1982), cited by part, paragraph and page. Works of Descartes cite two editions: CSM = *The Philosophical Works of Descartes* I–II, eds. Cottingham, Stoothoff, Murdoch (Cambridge: Cambridge University Press, 1984–1991) and AT = *Oeuvres de Descartes* I–XI, eds. Charles Adam and Paul Tannery (Paris: J. Vrin, 1996), cited by volume and page.

[2] Other questions raised by commentators about the opening moves of Book II include why Hume begins with the "indirect passions"—those passions that have objects, and why he starts with an account of "pride" as exemplary of such indirect passions. See, e.g., MacIntyre (2000), "Hume's Passions: Direct and Indirect," *Hume Studies* 26: 77–86, and Inoue (2003), "The Origin of the Indirect Passions in the *Treatise*: an Analogy Between Books 1 and 2," *Hume Studies* 29: 205–221.

[3] T I.iv.6, 251–252.

[4] T I.iv.6, 253.

J. Miller (ed.), *Topics in Early Modern Philosophy of Mind*, Studies in the History of Philosophy of Mind 9, 223–240.

"self" as "that succession of related ideas and impressions, of which we have an intimate memory and consciousness."[5] Perhaps then there is some way to reconcile the earlier skeptical account with Book II's treatment of the self.[6]

But a deeper puzzle remains. For what can Hume mean by saying that passions have *objects* at all? His account introduces a notion of "object" seemingly out of the blue. As we will see (Section 2), Hume takes pains to distinguish the object of the passion from other, related notions, such as its cause. For this reason, it seems likely that he uses "object" in the strong sense of classical intentionality: an intentional object is what the perception is about, what it is directed towards.[7] But that is just what makes the claim perplexing: little before this point in the *Treatise* showed any special concern to defend the intentionality of our mental states. If anything, Book I worked tirelessly to deflate our claims to having genuine objects in various important cases, e.g., when thinking of substance, of bodies, or even of self. Indeed, we may well wonder whether there is any place for a robust notion of intentionality in Hume's repertoire of impressions and ideas.

I think that there is a genuine, *prima facie* problem in finding a place for intentionality in the mechanics of mind that Hume presented in Book I of the *Treatise*—at least enough of one that the opening moves of Book II should come as a surprise. And I suggest that that trouble with intentionality is a legacy Hume inherits from his British Empiricist predecessors. That evaluation may seem obvious to some readers,[8] outrageous to others. So let me at least say a few words in explanation, for I do not mean to tar all early modern proponents of the "way of ideas" with the same anti-intentional brush. That view has indeed been propagated in some quarters, for instance, by the later Husserl, who takes Humean "sensationalist" "fictionalism" to be the culmination of Lockean, and ultimately Cartesian starting points. In this vein, he reads Hume's skepticism as representing a profound crisis for "objectivity" as such, one rooted in a sharp split between the ego and its intentional activity, so that it becomes impossible ever to "reach beyond the immanent sphere" and consciousness itself turns into a bit of an enigma.[9]

---

[5] T II.i.2, 277.

[6] For a reading that particularly contrasts between the practical bent of Book II and the preceding theoretical stance of Book I, see Purviance (1997), "The Moral Self and the Indirect Passions," *Hume Studies* 23: 195–212. See also Talia Mae Bettcher's essay in this volume.

[7] For this "classical" sense of intentionality, see Franz Brentano's famous characterization of intentionality as "what we might call, though not wholly unambiguously, reference to a content, direction toward an object (which is not to be understood here as meaning a thing), or immanent objectivity. Every mental phenomenon includes something as object within itself …."(*Psychology from an Empirical Standpoint*, transl. A.C. Rancurello, D.B. Terrell, and L. McAlister, London: Routledge, 1995), p. 88. In citing this passage, I don't mean to endorse any of Brentano's metaphysical commitments.

[8] One example is Barry Stroud who casually asserts that the "theory of mind [Hume] uncritically inherited leaves no room for what has been called the 'intentional' character of thought." See *Hume* (London: Routledge, 1977), p. 74.

[9] See *The Crisis of the European Sciences and Transcendental Phenomenology*, transl. D. Carr (Evanston: Northwestern University Press, 1970), pp. 82–90. To understand the full extent of

This seems to me to be too broad for plausibility, and to miss how much machinery the empiricists' continental counterparts had at their command for dealing with intentionality. Descartes, for one, does not so much insist that ideas *are* objects as that they *have* objects.[10] More obviously, Leibniz can distinguish between the rumble of *petites perceptions* and levels of increasing clarity and distinctness that characterize the objects of genuine consciousness and self-consciousness. These early continentals tend to use "idea" much more selectively than their British counterparts. And they typically suppose that an idea has at least enough internal structure to allow us to distinguish acts from objects, the qualities of the idea from its contents, and ideational contents in general from objective contents in particular. This gives room for innate ideas to play their role as *enabling* ideas, serving, for example, to structure sense-perceptions in a way that produces genuine objectivity in perception.[11]

I would suggest that we should look for any trouble Hume faces with intentionality a bit closer to home, particularly in the moves made by Locke and Berkeley to identify objects with ideas (or more broadly, with perceptions[12]), to reduce ideas to their contents, and to couple them with a resemblance theory of representation. The opening gambit here is Locke's claim that "idea is the object of thinking."[13] Despite its use of a seemingly intentional sense of "object," the claim works to eliminate any distinctive sense of the mind's being directed at a target object. For Locke assumes that ideas are what are present to mind, and presence to mind is *all* that is required to be an object. To be sure, he does allow that there may be additional kinds of objects, insofar as he sometimes specifies

---

Husserl's reading of Hume, we need to look ahead to later passages in the *Crisis*, where, e.g., Husserl takes the Hobbesian–Lockean (i.e., empiricist) tradition to require that "the soul is set off by itself in the closed unity of a space of consciousness," and to cause Berkeley and Hume to "press toward an immanent idealism which swallows up [the psychic counterpart to physical causal explanations,]" (p. 231) so that, ultimately, "the being of the world is incomprehensible." (p. 262) But he maintains it is Descartes who sets the whole process in motion (p. 83).

[10] Perhaps the places where Descartes comes closest to identifying ideas and objects are in the Replies appended to the *Meditations on First Philosophy*. For instance, the "geometric" account offered at the end of the Second Replies to the *Meditations on First Philosophy* tells us that "I understand ["idea"] to mean the form of any given thought, immediate perception of which makes me aware of the thought." (AT VII.160–161, CSM II.113) The Third Replies states simply that "I am taking the word 'idea' to refer to whatever is immediately perceived by the mind." (AT VII.181, CSM II.127) But if these are the best examples of such usage, they are not very good: *both* passages identify ideas with the "forms" of thought (e.g., the natural gloss on "perception which" in the first passage is to modify "*illam formam*"), and both suggest a contrast between form and object, or "what is represented by an idea."

[11] See, for instance, my gloss on the "wax passage" in Descartes's *Meditations* in Schmitter (2000), "The Wax and I: Perceptibility and Modality in the Second Meditation," *Archiv für Geschichte der Philosophie* 82: 178–201.

[12] There is some difficulty in deciding which term to use here, since Hume uses "idea" more narrowly than either Locke or Berkeley. "Perception" is the more neutral term, but it loses some of the flavour of the terms of art used in and about early modern philosophy. So until I turn specifically to Hume, I will use "idea" as interchangeable with "perception."

[13] E II.i.1, 104.

only that ideas are the *immediate* objects of thinking.[14] But adopting such a "contact" theory of the immediate intentionality of ideas seems to embrace too much: were mere contact with the mind sufficient for objectivity, we would lose any sense in which the mind can be *particularly* directed at a target object.

This threat is exacerbated by several other developments in the empiricists' increasingly atomistic views of the composition of ideas, which shape their understanding of the possible contents of ideas. In describing Locke's and Berkeley's views as atomistic, I do not mean anything particularly strong; I am not assuming that either philosopher believes that we receive ideas that are already divided into simples, or that we can actually divide complex ideas into purely simple ones.[15] All I mean is that they both hold that ideas can be analyzed without remainder into their simple component parts or the arrangement of those parts.[16] Any intra-ideational relations will be reducible either to those parts or to their arrangements. This in turn has the effect of reducing ideas largely to their contents, since there are no essential relations other than those holding between the simple parts. And that means that there is a very sharp distinction between the acts of the mind and its objects; ideas are nothing but the contents on which various acts are performed.[17] On the other hand, for that reason, there is no real distinction to be made between the contents and the objects of ideas: they are both simply identified with the idea proper. The upshot is an even more strictly stripped-down account of the objectivity of an idea: it simply provides content for some mental act.

There might still be recourse for a robust notion of intentionality, were the content provided by an idea-object itself able to direct the mind onto a further object. That is a possibility countenanced by Locke, who grants that ideas refer to, or represent things not present to the mind, which become objects for it by the intervention of the intentional operation of ideas. Locke explains this intervention by locating ideas within a doctrine of signs, and assigning the "power" to perceive the signification of signs to the act of perception.[18] But Berkeley's distinctive conception of ideas allows no such operation for ideas. Consider here how Berkeley argues against the claim that ideas could represent anything external on the basis that an idea can represent some X only insofar as the idea resembles, or is a copy of that X.[19] Berkeley assumes that the only relation of representation available is resemblance—a point that later figures centrally in his argument against the existence of material objects. It does not *ipso facto* rule out ideas having further objects. But it does betray a conception of what ideas are like that

[14] See, e.g., E IV.i.1, 525.

[15] In "Decompounded Complexity in Locke's Abstract Ideas," presented at the 2007 meetings of the Pacific Northwest–Western Canada Seminar on Early Modern Philosophy, D. Kenneth Brown plausibly argued that performing the act of "decompounding" on a Lockean complex idea does not merely resolve it into its component simple ideas.

[16] See, e.g., E II.ii.1, 119 or in a different vein, PHK I.124, 73.

[17] See, e.g., E II.i.22, 117 or II.xii.1, 163–164.

[18] See, e.g., E II.xxi.5, 236.

[19] PHK I.8, 25–26.

reduces them to their contents, *and* supposes that those contents have no properties different in kind from the properties of ordinary, particular independent existents (given that no such existents are material substances). Thus, there is no special intentional property in the idea of yellow; the idea simply *is* yellow, and as such, it can only bear the sorts of relations that yellow things bear (e.g., resemblance). In similar fashion, the idea of the species "dog" has no special abstract content; its content is just like that of any particular dog. Of course, since there are no material objects, that content is more like a fully determinate "picture" of a dog than like what we might previously have thought of as a material substance. But Berkeley advances a picture theory of ideas, not because he wants to capture some sense of intentionality found particularly in pictures, but because he conceives of pictures as ordinary things without matter. And just as everything becomes a mental entity for Berkeley, he loses the strong sense of intentionality that might seem the distinctive mark of the mental.[20]

The result of these several moves, I argue, is to flatten out the contents of ideas, so that they are all equally present to the mind that acts on them. Ideas press so closely on the mind, as it were, that there is no room left for talking about directedness, and no nuances in objectivity. There is certainly little room for admitting degrees of attention:[21] whatever is present to the mind is wholly present, and just is what it is, for all it may have causal connections with things in or outside the mind. In turn, presence to mind so sharply demarcates the contents of ideas from everything else that ideas lose any relation to whatever might serve as objects in an alternative sense. This thought is made explicit by Hume when he asserts that "the reference of the idea to an object [is] an extraneous denomination, of which in itself it bears no mark or character."[22] We are then left with only two options for explaining intentionality: being an object for the mind can be identified with mere presence, so that *all* perceptions become intentional objects,[23] or some sort of reference or representation can be established through "extraneous" relations, typically some form of resemblance, or copy-relation. Neither seems satisfactory for a robust, classical notion of intentionality, whereby the mind is genuinely *directed* upon a target-object. But they may be all that is available to Hume.

I do not doubt that Hume would think that much of what is worthwhile under the banner of intentionality can be captured either by the notion of perceptual contact, or by a suitably qualified copy-relation. The first gives us what it is to be a mental object; the latter does at least some explanatory work in accounting for how a perception comes to represent, or to have some content.[24] Hume uses

---

[20] Thanks are due to Talia Bettcher for pressing me to think more about the role Berkeley might play here.

[21] See E I.i.5, 45–46.

[22] T I.i.7, 20.

[23] See Annette Baier, *A Progress of Sentiments* (Cambridge, MA: Harvard University Press, 1991), pp. 116–118.

[24] I thank both Don Garrett and David Raynor for pushing me on this point.

"object" in both contexts: perceptions in general are objects,[25] and ideas, as copies, "always represent their objects or impressions."[26] The representation of objects by ideas may be the most familiar candidate for an account of intentionality in Hume;[27] to take it as such requires identifying intentionality with representation, and turning the account toward explaining how relations of representation are established. On that score, there seems little question that Hume makes the copy relation do a lot of heavy lifting, although it also seems implausible to hold that it is *sufficient* for constituting a relation of representation to some object.[28] But whatever story might be told there will not be much use for the curious treatment of pride, or indeed of any of the "indirect" passions; they may have objects, but they do not stand in any sort of copy-relation to those objects, and as we shall see further below, they do not "represent" those objects. Fortunately, we can find *another* candidate account of intentionality in Hume: one that offers a credible sense of how the mind can be directed toward an object by adopting a *holistic* approach to explaining how the passions shape our thought.

Stressing Hume's holism is not itself a particularly novel approach—at least not among contemporary Hume interpreters[29]—but I am not at all convinced that it has been exhausted. Even less explored may be how Hume addresses the nature of *in*tentionality by way of *at*tention. Attention, I maintain, was an issue absolutely central to early modern philosophy. It figured crucially in Leibniz's account of perception and apperception. It was important even in such seemingly disparate fields as that of method.[30] Most importantly for our purposes, attention and the passions were closely linked. Descartes, for instance, came to find the passions a

[25] T I.iv.5, 241.

[26] T I.iii,14, 157.

[27] See, e.g., Alexander Rosenberg, "Hume and the Philosophy of Science," in *The Cambridge Companion to Hume* (Cambridge: Cambridge University Press, 1993), pp. 65–66.

[28] As many have pointed out, neither resemblance, nor mere causation, nor both together seem to make the copy *about* anything: a bacterium is hardly *about* its parent-bacterium, even if it is a perfect genetic copy, any more than an etching produced from a master-plate is about that master. Still, Hume does put a great deal of weight on the copy-relation. For instance, he seems to find it a compelling enough account of representation to use it as a standard for ideas: without a corresponding impression from which it was derived (directly or "obliquely"), an idea lacks content—and indeed is no idea at all. See, e.g., what he says about our non-idea of "substance" (T I.i.6, 15–17). For a further discussion that locates the copy relation largely within an account of functional role, see Garrett (2006), "Hume's Naturalistic Theory of Representation," *Synthese* 152 (3): 301–319, especially pp. 307–313.

[29] This seems to be the approach of Amèlie Oksenberg Rorty, for instance, who stresses that nothing in our sense-impressions taken *seriatem*, nor in their contents, seems to give us any reference to the self. See Rorty (1982), "From Passions to Emotions and Sentiments," *Philosophy* 57: 167.

[30] For instance, Descartes' earliest philosophical work, the *Rules for the Direction of the Mind*, is chock-full of tips for developing our capacity for sustained and focused attention, as well as providing various systems of notation and symbolization to aid in this task. Many features of these systems are borrowed from earlier memory arts, but in Descartes's hands, they serve as a means for focusing present attention, so that we do not have to rely on memory to retain some content.

fruitful area for examining how we can direct and focus attention, as is particularly clear in the account of wonder [*admiration*] offered in *The Passions of the Soul*. Wonder is nothing but our passionate response to something we find worthy of attention. And under the rubric of wonder, Descartes explored issues ranging from the physiological conditions for paying attention to the natural rhythms by which attention is provoked, developed and eventually exhausted in the investigation of some object.

Hume, I suggest, is part of a thriving tradition that finds our passions a rich source of material for investigating the nature of attention. Yet he may go farther than his forerunners did. For Hume wants to show not just how we can pay attention to some object, but how we can use attention to generate robust objectivity for that object in the first place. That is a tall order, and I don't hope to satisfy it fully here. My aim rather is to consider how Hume sets out to explain the sort of intentionality the passions display, and how relations of association—which taken individually offer only causal connections—might help to produce genuine intentionality. My main example will be the passion of pride. There's a reason for that, but I trust that much of what I will say will be readily generalizable to other indirect passions, and perhaps to some of the direct as well.

2 Hume's Peculiar Account of Pride

The difficulty in explaining intentionality is not just a matter of the sort of machinery Hume receives from his predecessors. Hume himself seems deliberately to make trouble. Consider the infamous passage at T 415, which specifically addresses the intentionality of the passions:

> A passion is an original existence, or if you will, modification of existence, and contains not any representative quality, which renders it a copy of any other existence or modification. When I am angry, I am actually possesst with the passion, and in that emotion have no more a reference to any other object, than when I am thirsty, or sick, or more than five foot high.[31]

Terence Penelhum considers this passage to constitute a "wildly implausible denial of the intentionality of passions and desires."[32] And Annette Baier refers to it as "that unfortunate paragraph," which gave Hume ample occasion to "repent his 'Haste' in this 'defective' passage."[33] It is hard not to sympathize with these responses: to the already daunting task of figuring out how passions can have objects, we here seem to face an explicit denial that they do. Of course, some features of the passage are fairly tractable: for instance, there are several different senses of "original," and whatever Hume means here, it is surely not that the

[31] T II.iii.3, 415.

[32] "Hume's Moral Psychology," in *Cambridge Companion to Hume* (Cambridge: Cambridge University Press, 1993), p. 128.

[33] *A Progress of Sentiments* (Cambridge, MA: Harvard University Press, 1994), p. 164.

passions fail to count as impressions of *reflexion*, that is, as impressions following upon certain ideas. Instead, the denial of "any representative quality" seems entailed simply by the status of passions as *impressions*. In contrast to an idea—be it of anger, or of an infuriating person, or of anything else—the passion is no copy, but something else altogether.

None of that, however, rules out holding that the passion is directed at the schmuck whose misdeeds prompt anger. The schmuck may be the cause of the anger, but insofar as the anger is particularly directed at him, he is also the object. The anger is somehow *about* that object in a way that is characteristic of the passion and serves to differentiate it from a host of similarly aversive affects. There is no reason to think that the "unfortunate paragraph" raises a problem for this sort of directedness, unless the "representative quality" provided by the copy relation—the kind of relation that ideas bear to their source impressions—were the sole means for establishing aboutness. More difficult is Hume's claim that when "possest with the passion, I make no reference to any other object." Perhaps this is simply a dramatic version of the previous claims, or perhaps Hume means something distinctive by "reference to any other object." I have no ready answer.[34] However we read them, these difficult passages set us the task of somehow reconciling the more-or-less self-contained affective character of a passion with its intentionality. In that light, we might distinguish here between representation and intentionality: a passion represents nothing, yet still may have an intentional object.[35]

That is the case with the indirect passion of pride. It is the object, self, that most markedly distinguishes pride from other, similarly pleasant passions, such as love. This object is to be distinguished from the cause of the passion, whether it is a matter of the quality that directly causes the passion, or of the "subject" in which that quality inheres.[36] Indeed it is the necessity of its object that sets pride among the indirect passions, in which the idea of the object of the passion follows upon the distinctive impression of reflexion, which in turn is layered on the idea of the cause of the passion, aroused by an original impression:

> Pride and humility, being once rais'd, immediately turn our attention to ourself, and regard that as their ultimate and final object; but there is something farther requisite in order to raise them …. The first idea, that is presented to the mind, is that of the cause or productive principle. This excites the passion, connected with it; and that passion, when excited, turns our view to another idea, which is that of self. Here then is a passion plac'd betwixt two ideas, of which the one produces it, and the other is produc'd by it. The first idea, therefore, represents the *cause*, the second the *object* of the passion.[37]

---

[34] In discussion, Anna Stoklosa made the interesting suggestion that Hume might here refer specifically to Descartes's description of how we "refer" passions to the soul in the *Passions of the Soul*; see, e.g., AT XI.349, CSM I.338. This explains the denial more plausibly than would, say, glossing it as implying some contemporary theory that takes reference to entail an existential claim. But it would seem a somewhat cryptic allusion on Hume's part.

[35] Alison Simmons similarly distinguishes representation from intention in her paper. See p. 112. [Ed. note.]

[36] See T II.i.2, 279.

[37] T II.i.2, 278.

Yet Hume's declared reasons for requiring a distinction between the status of the ideas surrounding pride are curious: the latter idea, that of self, cannot be counted the cause of the passion, for it is not "sufficient alone to excite" the passion. Since humility shares the same object, and humility and pride are opposites, neither passion could ever arise, were the idea of self alone the cause. Similarly, the same cause—the pleasure-inducing quality of some subject—can also produce the passion of love when the idea that follows is not self, but another. It is to differentiate pride from these other passions that Hume declares that we must "make a distinction." The distinction that Hume proceeds to make, of course, is that between cause and object: since self is not the cause, it must be the object "that to which [the passions of pride or humility] direct their view, when excited."[38] But this is quite a leap: why should we draw the distinction so? There seems any number of other possibilities that have not yet been eliminated: one might, for instance, hold that its affective character is sufficient to distinguish pride, or perhaps require its cause to be specified more finely (e.g., as a complex including a pleasurable impression involving a relation to self). True, this latter possibility could violate the second rule "by which to judge of causes and effects," which Hume laid down in Book I,[39] if the idea of self does not appear prior to the passion of pride.[40] And Hume seems reluctant to depend heavily on the affective character of various passions—the sense in which they are "simple and uniform impressions"—preferring instead "to pretend to a description of them, by an enumeration of such circumstances, as attend them."[41] Nonetheless, making the distinction by introducing a notion of object seems an unwarranted and excessive move on Hume's part.

That the introduction of an object does exceed what is required by Hume's arguments may, however, indicate several important points: first, that we are dealing with a genuinely intentional notion of object, and second, that it is not reducible to the other relations, causal and associative, in play at this point of the *Treatise*. And it does seem that Hume needs this excessive notion of objectivity to distinguish pride from, on the one hand, mere pleasure, and on the other hand, from the other passions that can follow pleasurable impressions. Pride may indeed have a distinctive affective component, a special feel, but Hume would have to count that "a simple and uniform impression," for which "tis impossible we can ever, by a multitude of words, give a just definition."[42] Perhaps Hume assumes either that that feel is insufficient to describe genuine pride, or perhaps he thinks that the affect would not be the distinctive affect of pride in the absence of its object. Or perhaps he simply wants to describe the typical conditions in which pride appears. These possibilities raise difficult questions about how to understand the identity conditions for various impressions, but we do not need to decide them here. All we need to note is that Hume's introduction of the object of pride is

[38] T II.i.2, 278.

[39] T I.iii.15, 173.

[40] However, I will argue below that this idea need not be considered to appear after pride in time.

[41] T II.i.1, 277.

[42] *Ibid*.

crucial to his account of the passion. Although it is not a copy, "representation," description, or even effect of the self, pride is still somehow typically *about* the self. A warm glow that is not about the self would probably be unrecognizable as pride. But at this point in the *Treatise*, Hume may lack the resources needed to explain this aboutness.

What do we come to learn about the connections between pride and the idea of self? Subsequent sections analyze a whole host of relations of association that the idea of self bears to the various other ideas and impressions surrounding the passion of pride. These relations are causal. What they cause are not ideas or impressions as such, but rather their occurrences, their becoming present to mind. As such, their explanatory role lies not in a genetic account of how we come to have various mental contents in the first place, but in an account of the structure and train of perceptions accompanying the occurrence of the passion. Hume first describes the relations involved in pride by locating the impression of reflection between two ideas, "of which the one produces it, and the other is produc'd by it."[43] Here he inverts the causal relation each idea bears to the passion in order to distinguish between cause and object. However, he goes on to complicate the account of the relations in which the passion and its object are embedded. Instead of a simple sandwich, he ends up with a four-fold picture in which two impressions are related by resemblance, and the whole shows the "double relation of ideas and impressions."[44] In the case of pride, the two impressions are simple pleasure and the impression of reflection specific to pride; the two ideas are the idea of the cause and the idea of the self. Such a double relation characterizes all the indirect passions, and differences in its structure allow us to capture what is typical of different passions:

> Upon the whole, pride is connected with humility, love with hatred, by their objects or ideas: Pride with love, humility with hatred, by their sensations or impressions …. [N]othing can produce any of these passions without bearing it a double relation, *viz*. of ideas to the object of the passion, and of sensation to the passion itself.[45]

In the double relation typifying pride, the idea of self no longer appears to be "produced" simply by the passion of pride. Rather, it is associatively related to—and thus prompted by—the idea of the subject. That subject also has some quality capable of exciting an impression of pleasure; and pleasure leads to pride by way of a relation of resemblance. Consider, for example, the pleasurable passion of pride Mr. Darcy may take in his beautiful estate:

	Cause (*quality*)	*Association*	*Passion*
Impressions:	Pleasure	(resemblance)→	Pride
	↑ ↑		↓ ↓
Ideas:	My Estate	(causation)→	Self (Mr. Darcy)
	Cause (*subject*)	*Association*	*Object*

[43] T II.i.2, 278.

[44] T II.i.5, 286.

[45] T II.ii.2, 333.

The idea of self is thus connected in multiple ways with the impression of pride, but there is no presumption that it appears after the passion. If anything, it should appear after the idea of the subject, to which it is specifically linked by relations of association, and thus to be more or less contemporaneous with the impression of pride.

Nonetheless, Hume does take it that the whole double relation "concurs" in the idea of self. Like all associations, the double relation gives regularity to the "changeableness" and inconstancy of our thoughts.[46] But the double relation does more: it provides a "double impulse," so that "these two kinds of association … very much assist and forward each other, and … the transition is more easily made where they both concur in the same object."[47] Relations of association provide an impulse and direction to thought, the force of which is doubled when two relations have the same direction. In the double relation involved in pride, that direction is provided by the idea of self. Thus not only does the double relation impose more than usual shape on the chaotic stream of our thoughts, it furnishes a natural resting place for our thoughts in the idea of self.

Hume's account of the relation between pride and self has several other features worth noting. First is the sheer diversity of subjects capable of rousing pride: "Upon my consulting experience, … I immediately find a hundred different causes, that produce pride …."[48] The causes of pride are "natural," but not "original;" almost anything can be a cause of pride, as long as it possesses a pleasure-inducing quality and bears a sufficient relation to self. Second, it seems that almost any pleasurable passion can readily be converted into pride by way of resemblance if the rest of the double relation is in place. That I take it is the lesson of the operation of sympathy in our "Love of Fame." Hume here tries to explain why we seek the good opinion of others—something we value for its own sake, independently of any advantages that might accrue to us thereby. Through the vivacity-increasing mechanism of sympathy, we find the approbation of others genuinely pleasurable, since it is a response to our praiseworthy qualities—which located in others prompt love—but in us, provide an occasion for pride. More generally, the operation of sympathy allows us to convert ideas of different, yet resembling passions into a genuine passion of pride. Because of sympathy our virtues too can be a cause of pride.[49] What is needed is that the progress of

[46] T II.i.4, 283.

[47] T II.i.4, 283–284.

[48] T II.i.5, 288.

[49] Hume recognizes virtue as a source of pride in Book II (T II.i.7, 297), but we need to look to the machinery of Book III to see the conversion to pride. Book III's account of our ability to recognize virtues, especially artificial virtues, requires both the operation of general rules and the development of "a general point of view" in order to feel the pleasurable impressions that are the

perceptions—of impressions of sensation to passions, and of idea to idea—provides sufficient vivacity and pleasure in a direction that culminates in the idea of self. It is because the progress of our perceptions, in tandem with the operation of sympathy and the influence of general rules, so often leads our thought back to the idea of self that the causes of pride proliferate. The double relation in which these causes figure may be doubled and redoubled again; what counts is the nature and direction of the impulses therein provided. There are, of course, "limitations to this system," that is, cases where "agreeable objects, related to ourselves, by an association of ideas, and of impressions" do *not* produce pride. For instance, merely being present at a glorious feast does not make us proud, however delightful the feast may be; only the host feels pride.[50] But these are limitations on how readily the association of our ideas leads to the idea of self. Not everything that possesses some pleasurable quality and is somehow related to self succeeds in directing our thought to the idea of self. And so they fail to be causes of pride.

3 The Train of Perceptions and the Self as Focal Point

At T 385, Hume tells us that "'tis not the present sensation or momentary pain or pleasure, which determines the character of any passion, but the general bent or tendency of it from the beginning to the end."[51] This is a maxim I think we should take to heart for understanding the intentionality of the passions. Just as determining the "character" of a passion requires a holistic view, so too does understanding its intentionality. For as we have seen, in the tradition Hume inherits, single perceptions lack the sort of internal structure that might produce a reference to, or directedness on an object. In fact, though, Hume's thought here seems to pull in two different directions: one atomistic, and one much less so. These two different directions suggest two different ways for drawing boundaries around the impression of pride. On the one hand, pride might be identified only with a single perception (i.e., that on the upper right hand corner of the chart above). As such, the passion maintains a causal association with the idea of self. But this connection cannot be a necessary one, in light of Hume's allegiance to a principle enunciated early in Book I: "*that all ideas, which are different, are separable*."[52] Considered merely as the momentary, or "simple and uniform" impression, pride can be separated from its object, or indeed from the whole double relation, and still remain pride. This may be the most natural way to read Hume, identifying pride strictly with the single impression in isolation. On

origin of all our moral sentiments (T III.ii.6, 531; III.iii.1, 581–582). These too require the vivification provided by sympathy and a re-orientation of the direction of our passions, and only so can virtue be a cause of pride.

[50] T II.i.6, 290.

[51] T II.ii.9, 384–385. The context of the remark is his discussion of how either benevolence or contempt might arise from pity, but the point seems perfectly general.

[52] T I.i.7, 24; see also I.i.3, 10.

the other hand, doing so means that we will fail to capture the "character" of pride. This is the less atomistic side of Hume's views: whatever pride is "like" requires looking a bit beyond the single impression to the whole double relation and the train of perceptions surrounding the passion of pride. Perhaps this "character" does not provide precise identity conditions for the impression, but fortunately, we do not have to decide what those might be. At the very least, the holistic character describes the typical context in which the passion appears, and without which it may be unrecognizable.

So let us look beyond the narrow boundaries of the single perception to the broad character that comprises the object of an indirect passion for an account of its intentionality. Now, that object is an idea, and that idea is linked through relations of association to the other impressions and ideas characterizing the passion. But those relations are not directionless. The effect of the relations of association is to turn and return our attention to the object-idea. The associative relations themselves are causal, that is, they describe how one perception gives rise to another perception. But they trace a pattern that is more than merely causal. Considered holistically, the passion is characterized by a structure in which the object appears as a focal point for the mind, indeed a central point around which an entire train of perceptions revolves. This, I suggest, is the sense in which a passion is *about* its object. As Hume tells us in the case of pride, the passion "turn[s] our attention," or "turns our view" to the object; the object is where "the view always fixes."[53] The account gives us, I think, a genuine notion of intentionality, of what it is for a passion to be about an object. But it may not be one that can be cashed out by the familiar means of propositional attitudes[54]—for it need involve no propositions—nor perhaps by familiar notions of representation. The passion is about the object, because it draws our attention to it. The object is an object of attention, not the subject of a proposition, nor the reference of a representation. The mechanism by which attention is focused on this object is causal; the result, however, is intentional. It is for this reason that I think it ultimately makes good sense for Hume to say that pride "produces" the idea of self: it is not because the impression of reflection *alone* prompts the idea of self by means of relations of association, but because the entire passion focuses attention on the idea of self. And in so doing, it puts some flesh on the bones of what may begin as quite a skeletal idea. Indeed, "the" idea of self may not be exactly the same in all these cases, a possibility that would accord well with Hume's denial that the perception of self is something "simple," much less "continu'd."[55] All that is needed is that closely resembling ideas play the same role within the repeated pattern characterizing the passion. This would give the self a "fictitious" identity,[56] which title is not meant to deny its reality, but to indicate that the identity is constructed

[53] T II.i.2, 277–278.

[54] For a different view, see the famous article by Davidson (1976), "Hume's Cognitive Account of Pride," *Journal of Philosophy* 73: 744–757; and for criticisms on rather different grounds than mine, see Baier (1978), "Hume's Analysis of Pride," *Journal of Philosophy* 75: 27–40.

[55] T I.iv.6, 252.

[56] T I.iv.6, 259.

by patterns of the imagination, or of the mind more generally.[57] Indeed, this self may be a "fiction" in a double sense: "produced" by pride and forged into some sort of unity by the imagination.

Now, if I am right, not only does the intentional structure of the passion fail to be found internally to any single perception, it is not synchronically complete. Rather it is a temporally extended structure, one produced by a train of perceptions considered in its "general bent or tendency … from beginning to end." For this reason, the self that is the object of pride is a far cry from the "I think" that accompanies all of our representations, even if that could become the object of a representation, and even if it were nothing but the bundle of representations. Rather, the self that is the object of pride should be understood as a full-blooded self—a self outfitted with its qualities, possessions, relations, likes and dislikes—it is a character, or personality.[58] As we saw in Section 2, Hume tells us that almost anything can serve as a subject for pride, as long as it has pleasurable qualities and a relation to self. At least part of his reason for casting his nets so widely here is, I think, to allow for the sheer variety of features that may make up a particular personality. And if it is such a full-blooded personality that is the object of the passion of pride, it makes sense that its intentionality should be temporally extended—both because personalities are temporally extended, and because the appreciation of personalities takes time.

The analysis of intentionality that looks not to the contents of individual perceptions, but to the structure of a train of perceptions shares a good deal with Book I's treatment of the objects of ideas such as personal identity, or substance. The machinery is pretty much the same, although the tasks and results differ enormously. Book I does indeed aim to deflate our pretensions to have a genuine object in mind when talking about, e.g., substance, showing instead that all we have is a collection of perceptions connected by relations of association. But, of course, that is all we have with the passions (with even fewer relations of associations available to link the impressions). And Book I showed how these relations could serve to structure collections of ideas into objects of a sort, using fictions of varying degrees of ingenuity. What gives Book I much of its skeptical cast, however, is the danger that these might turn out to be nothing but *mere* fictions—slips and prejudices. If so, then our talk of an "object" would be little more than an unnoticed error, picking out no real object-idea on which the

[57] Drawing from T 11, Don Garrett suggests that "all fictions involve proceeding as though we have an idea representing something when in fact we do not have an idea that can properly represent that thing, because we lack an idea derived from it." (Garrett 2006, p. 313, n. 17) Perhaps, but that may not fully capture the sense in which a fiction is a "construction" of the imagination (in an echo of the Latin root "facere," to make, or construct). Both senses may explain why Hume often takes "fictions" to accompany mistakes (see, e.g., T I.iv.6, 255) and why calling something a fiction sometimes constitutes a criticism of the attendant beliefs. But even if fictions are "imaginary," Hume does not hold that they are inevitably false (see, e.g., the use at T I.iv.2, 201). For further discussion, see Baier (1994), *A Progress of Sentiments*, p. 103 (on the negative point), and for a particularly strong view, Traiger (1987), "Impressions, Ideas and Fictions," *Hume Studies* 13: 381–399.

[58] Rorty (1982) makes a similar point in "From Passions to Emotions and Sentiments," p. 168.

relations of association allow us to focus. Hume's holistic account of the passions' intentionality does not allow us to create an object wholesale, where no idea previously exists, or where no idea has the sorts of associative relations that will be the glue for the whole structure. What it does allow is that an idea may become an object, and that a previously thin, faint idea (e.g., the idea of self) may become more full-blooded as we turn our attention to it again and again. That can happen when something is a subject of passionate concern, but perhaps not as readily when we restrict ourselves to theoretically conceived ideas alone.

Still, there is at least one remaining problem with this sort of holistic approach that may threaten to turn the object into a delusion of a different kind. It tells us to look for the intentionality of the passions not in the individual perceptions, but in an entire train of perceptions. But where should that train come to an end? Relations of association provide the glue holding the structure of intentionality together, but they also appear everywhere in our thought. Presumably, whenever our thoughts show some sort of continuity, even merely causal continuity, relations of association are at work. Even if all we have is a succession of different thoughts of different objects, that succession is still associatively linked, unless it is interrupted by some overwhelmingly intrusive impression (e.g., a poke in the eye with a sharp stick), or deep sleep, or the like. Yet we don't take the sorts of progression of thoughts that happen when we daydream, or let our mind wander, or simply look around us at the world to have any specific intentional object. Of course, they lack a focal point, a center of attention, and that may be explained by saying that the meanderings of associations fail to describe any coherent structure. But suppose we *do* find our attention returning to some idea that we cannot escape, say the idea of an impending exam, or of an overdue bill, or what-have-you. In that case, we have a kind of center of attention, an *obsession*, which intrudes on our thoughts again and again. Yet in these cases, we don't suppose that the rest of our thoughts are *about* whatever it is we cannot help fixating on—or if we do, we do so because we are in the grip of an obsessive disorder. So we have a puzzle: our perceptions are typically linked together in an associative web, but associative links seem all that is available to determine particular perceptual structures within that web. How then will we differentiate an intentional structure sufficient to give us an object?

How we answer this question depends on how we go about distinguishing what belongs to the passion from the various other perceptions that bear (perhaps with drastic contingency) associative relations to it. As we have already seen, that it not an easy issue to settle, and I will remain steadfast in my refusal to try to determine strict Humean identity conditions for a passion.[59] But that shouldn't prevent us from further consideration of what might give the character of a passion its distinctive shape. Consider the way in which a passion is about its object. The passion of pride directs attention to the self, but it is not merely the perception of self. We think about ourselves in a pleasurable, satisfied, warm way—in short, with pride. This is a tempting place to locate the affective character

[59] I suspect that Hume's thought is simply too divided for there to be an unambiguous solution to this question.

of the passion. We might, for instance, say that pride *colours* our perception of its object, that it constitutes a "mood" through which we see the object of the passion. On a view of this sort, the train of perceptions that characterizes the passion is marked off from the rest of our thinking by its distinctive colouration. Pride, for instance, casts a glow over the various perceptions that it runs through; we might think that it is that glow that links them together, and thus, that it is through that glow that the idea of self emerges as an object. *Mutatis mutandi*, the mood of humility paints the self in unpleasant and uncomfortable colours.

But notions of affective colouration or mood seem alien to Hume—at best, anachronistic, at worst, at odds with the machinery of simple and complex perceptions and the principles of association on which Hume relies throughout the *Treatise*. But then how are we to account for the "character" of our passions at all? Well, we might look to the patterns that emerge from the association of perceptions; Hume uses such patterns in a number of places to explain how structure can be imposed on the often chaotic stream of our perceptions. Particularly important is the effect those patterns have on our experience of our own perceptions. Now, as we have already seen in the case of the indirect passions, the double relation gives us several, mutually reinforcing patterns. First are relations of resemblance holding between impressions; in the case of pride, the impressions each have a resembling pleasurable quality. Indeed, there is no reason not to extend that resemblance to the associated ideas: they will not share the "same" pleasure, but they will be similarly pleasant.[60] So, the train of perceptions characterizing pride might be linked simply by their resembling qualities. And the *effect* of that resemblance could be considered to impart a certain "colour" to the whole.

But that is surely not all: the train is linked by causal connections. Indeed, those connections were the very reason for considering the perceptions to form a train, whereby the appearance of the cause prompts the appearance of the passion and of its object. Moreover, these causal connections form a *regular* and *entrenched* pattern.[61] How regular a pattern must be to consider it entrenched is a matter of degree; there are other, less entrenched associative paths the mind can take, which still display enough regularity for causality. But that a pattern of association is entrenched makes a big difference in its effect on the mind. For one, it produces a habit of expectation, which will itself promote the regularity with which the train of perceptions appears. Such a habit will also generate resemblances between instances of the train of perceptions; they will in turn reinforce both the causal pattern, and the resemblances between the perceptions. Insofar as the relations of association between the perceptions characterizing pride are reinforced,

[60] Even at his most atomistic and nominalist, as when he considers the simple ideas of colours, Hume allows various degrees of resemblance to hold between perceptions. And the "feeling" of pleasure need not be restricted to impressions; ideas will have this quality in much fainter, but still resembling form. Indeed, ideas of pain and pleasure *must* still have a whiff of pain or pleasure about them.

[61] I owe this suggestion to Don Garrett.

and the transitions of the mind thereby eased, we will run the train of perceptions together as if they were a single whole, differentiated from whatever other perceptions are less regularly associated with the members of the train. Such running together of distinct perceptions is one of the most important moves Hume describes in his skeptical accounts of the construction of fictions such as personal identity. In this case, however, it seems innocuous, exciting no unwarranted beliefs in ontologically suspect objects. What it produces is merely the colouration, or mood that characterizes the passion in its "general bent or tendency." That is an *effect* of linking the train of perceptions together from beginning to end, however, not its cause.

My aim throughout this discussion has been to understand how the indirect passions illustrate an alternative conception of intentionality in general. But there is a reason why pride offers the foremost example of how a perception can be directed on an object. More than any other passion, pride shows both how a passion can direct our attention onto an object, and in doing so provide the structure to a train of passions that will give them a distinctive character. Our other indirect passions, and perhaps even our direct passions, may also serve to direct our attention to particular ideas that stand as their objects, or in the case of direct passions, their causes. But no other idea looms as large in our consciousness as our idea of self—particularly, the full-blooded idea of a self equipped with properties and property, qualities, relations and the like. Because "we are at all times intimately conscious of ourselves," our ideas of ourselves are particularly lively and vivacious. For this reason, almost all roads lead to the idea of self; it "becomes present to the mind on the smallest hint and most trivial relation," and once present, "it engages the attention, and keeps it from wandering to other objects, however strong may be their relation to our first object."[62] This does not mean that our idea of self is particularly distinct, accurate, or detailed—only that the idea of self exercises a greater gravitational pull over our attention than does any other perception. No other idea is as likely to be excited by the various relations of association our perceptions bear to each other, especially when it is forwarded by pleasurable impressions, and no other idea compels our attention as does the idea of self. Likewise, pride itself is easily provoked and readily maintained. Love of a relation typically converts into pride, but not vice versa: we tend to remain consumed with pride. All of this is part and parcel of the attention-grabbing features of the idea of self, and so, pride should show more clearly than any other passion what it is for a passion to have an object. Of course, humility too takes self as its object, and so one might think it will share the exemplary status of pride. But the crucial point of difference here is that humility is not only a painful passion, but also an enervating one. Pride, in contrast, is invigorating, and its invigorating character carries over to the way its object engages our attention. We

[62] T II.ii.2, 339.

have an object, because the causal structure of our minds finds a focal point, and we are so constituted as to return to it again and again. There is some slight irony in advancing this thoroughly naturalist tale about the mind through the character of pride. Hume, I suspect, was amused.[63]

[63]I'd like to thank several audiences who heard different versions of this paper, many of them even less polished than this one: first is the audience at the University of Alberta Philosophy Department Colloquium in 2004, particularly Edwin Etieyebo, who raised a number of important points in his thoughtful commentary. I also appreciate the help of the audience at the 2004 Hume Society Conference in Tokyo, particularly Haruko Inoue, who was an especially astute commentator; Rachel Cohen, Geoff Sayre-McCord and Hide Ishiguro contributed a number of insights. Also very helpful was the session on this project at the conference, "Topics in Early Modern Philosophy of Mind," ably organized by Jon Miller at Queen's University in November 2006; thanks are due especially to Talia Bettcher, Don Garrett, David Raynor, and Ana Stoklosa. I would also like to thank my hard-working Research Assistants, Yual Chiek and Alan McLuckie, who helped a great deal with research, writing and the correction of many a typo. Barbara S. Schmitter also deserves gratitude for catching a number of mistakes, both stylistic and substantial; however, to express that would break the habits of a lifetime, so I guess I won't do so. Support for this project was provided through both the Endowment Fund for the Future/Support for the Advancement of Scholarship at the University of Alberta, and by the Social Sciences and Humanities Research Council of Canada, for which I am grateful.

Bibliography

Ainslie, Donald. 2001. Hume's Reflection on the Identity and Simplicity of the Mind. *Philosophy and Phenomenological Research* 63 (3): 557–578.

Ainslie, Donald. (no date). Hume's Anti-*Cogito*. Unpublished manuscript.

Alanen, Lilli. 1994. Sensory Ideas, Objective Reality, and Material Falsity. In *Reason, Will, and Sensation: Studies in Descartes' Metaphysics*, ed. John Cottingham. Oxford: Clarendon Press.

Alanen, Lilli. 2003a. *Descartes' Concept of Mind*. Cambridge, MA: Harvard University Press.

Alanen, Lilli. 2003b. The Intentionality of Cartesian Emotions. In *Passion and Virtue in Descartes*, eds. Byron Williston and André Gombay. New York: Humanity Books/Prometheus Books.

Alquié, Ferdinand. 1974. *Le Cartésianisme de Malebranche.* Paris: Vrin.

Ameriks, Karl. 2000. *Kant's Theory of Mind*, 2nd edition. Oxford: Oxford University Press.

Aquila, Richard. 1974. Brentano, Descartes, and Hume on Awareness. *Philosophy and Phenomenological Research* 35 (2): 223–239.

Aquila, Richard. 1995. The Content of Cartesian Sensation and the Intermingling of Mind and Body. *History of Philosophy Quarterly* 12 (2): 209–226.

Arbini, Ronald. 1983. Did Descartes Have a Philosophical Theory of Sense Perception? *Journal of the History of Philosophy* 21 (3): 317–337.

Armogathe, Jean-Robert. 2000. The Rainbow: A Privileged Epistemological Model. In *Descartes' Natural Philosophy*, ed. Stephen Guakroger. London: Routledge.

Arnauld, Antoine. 1780. *Oeuvres de Messire Antoine Arnauld*. Paris: Sigismond D'Arnay & Compagnie.

Arnauld, Antoine. 1986. *Des Vrais et Fausses Iddées*. Paris: Fayard.

Arnauld, Antoine. 1990. *On True and False Ideas*, transl. Stephen Gaukroger. Manchester: Manchester University Press.

Atherton, Margaret. 1983. The Coherence of Berkeley's Theory of Mind. *Philosophy and Phenomenological Research* 43 (3): 389–399.

Atkins, Kim, ed. 2005. *Self and Subjectivity*. Oxford: Blackwell.

Attfield, Robin. 1977. Clarke, Collins and Compounds. *Journal of the History of Philosophy* 15 (1): 45–54.

Ayers, Michael. 1981. Mechanism, Superaddition, and the Proof of God's Existence in Locke's *Essay*. *The Philosophical Review* 90 (2): 210–251.

Ayers, Michael. 1983. Berkeley and Hume: A Question of Influence. In *Philosophy* and History, eds. Richard Rorty, Jerome B. Schneewind and Quentin Skinner. Cambridge: Cambridge University Press.

Ayers, Michael. 1998. Theories of Knowledge and Belief. In *The Cambridge History of Seventeenth Century Philosophy*, vol. 2, eds. Daniel Garber and Michael Ayers, 1003–1061. Cambridge: Cambridge University Press.

Baier, Annette. 1978. Hume's Analysis of Pride. *Journal of Philosophy* 75 (1): 27–40.

Baier, Annette. 1991. *A Progress of Sentiments*. Cambridge, MA: Harvard University Press.

Beardsley, William H. 2001. Berkeley on Spirit and Its Unity. *History of Philosophy Quarterly* 18 (3): 259–277.

Belfrage, Bertil. 1987. A New Approach to Berkeley's *Philosophical Notebooks*. In *Essays on the Philosophy of George Berkeley*, ed. Ernest Sosa. Dordecht: D. Reidel.

Bennett, Jonathan. 1971. *Locke, Berkeley, Hume: Central Themes*. Oxford: Clarendon Press.

Bennett, Jonathan. 1984. *A Study of Spinoza's Ethics*. Indianapolis, IN: Hackett.

Bennett, Jonathan. 1990. Spinoza and Teleology: A Reply to Curley. In *Spinoza: Issues and Directions. The Proceedings of the Chicago Spinoza Conference (1986)*, eds. Edwin Curley and Pierre-François Moreau. Leiden: E.J. Brill.

Bennett, Jonathan. 2001. *Learning from Six Philosophers*, 2 vols. Oxford: Oxford University Press.

Bennett, Jonathan and Alston, William. 1984. Identity and Cardinality: Geach on Frege. *The Rhilosophical Review* 93 (4): 553–568.

Berkeley, George. 1948–1957. *The Works of George Berkeley, Bishop of Cloyne*, eds. A.A. Luce and T.E. Jessop, 9 vols. London/Edinburgh: Thomas Nelson & Sons.

Berkeley, George. 1982. *A Treatise Concerning the Principles of Human Knowledge*. Indianapolis, IN: Hackett.

Bettcher, Talia Mae. 2007. *Berkeley's Philosophy of Spirit: Consciousness, Ontology and the Elusive Subject*. London: Continuum Press.

Bobro, Mark and Paul Lodge. 1998. Stepping Back Inside Leibniz's Mill. *The Monist* 81 (4): 554–573.

Bolton, Martha. 1986. Confused and Obscure Ideas of Sense. In *Essays on Descartes' Meditations*, ed. Amèlie Oksenberg Rorty, 389–403. Berkeley, CA: University of California Press.

Bouillier, Francisque. 1868. *Histoire de la Philosophie Cartésienne*. Paris: Ch. Delagrave et Cie.

Boyer, Carl. 1959. *The Rainbow*. Princeton, NJ: Princeton University Press.

Boyle, Robert. 1687. *Disquisition About the Final Causes in Nature*. London.

Boyle, Robert. 1688. *A Disquisition about the Final causes of Natural Things*. Printed by H.C. for John Taylor, St. Paul's, London.

Brentano, Franz. 1995. *Psychology from an Empirical Standpoint*, transl. A.C. Rancurello, D.B. Terrell, and L. McAlister. London: Routledge.

Broughton, Janet. 1986. Causes and Natural Change in Descartes's Philosophy. In *Human Nature and Natural Knowledge: Essays Presented to Marjorie Grene on the Occasion of Her Seventy-Fifth Birthday*, eds. Alan Donagan, Anthony N. Perovich Jr., and Michael V. Wedin. Dordrecht, the Netherlands: Reidel.

Brown, D. Kenneth. 2007. Decompounded Complexity in Locke's Abstract Ideas, presented at the 2007 meetings of the Pacific Northwest–Western Canada Seminar on Early Modern Philosophy.

Brown, Deborah. 2006. *Descartes and the Passionate Mind*. Cambridge: Cambridge University Press.

Butler, Ronald J, ed. 1972. *Cartesian Studies*. Oxford: Blackwell.

Butler, Samuel. 1970. *Characters*. Cleveland, OH: Press of Case Western Reserve University.

Carriero, John. 2005. Spinoza on Final Causality. In *Oxford Studies in Early Modern Philosophy*, vol. 2, eds. Daniel Garber and Steven Nadler, 105–147. New York: Oxford University Press.

Chalmers, David. 2003. Consciousness and Its Place in Nature. In *Blackwell Guide to Philosophy of Mind*, eds. Stephen Stich and Fritz Warfield. Oxford: Blackwell.

Chappell, Vere. 1986. The Theory of Ideas. In *Essays on Descartes' Meditations*, ed. Amèlie Oksenberg Rorty, 177–198. Berkeley, CA: University of California Press.

Chisholm, Roderick. 1969. On the Observability of the Self. *Philosophy and Phenomenological Research* 30 (1):7–21. Reprinted in *Self-Knowledge*, ed. Quassim Cassam, 94–108. Oxford University Press: Oxford, 1994.

Chomsky, Noam. 1995. *The Minimalist Program.* Cambridge, MA: MIT Press.

Chomsky, Noam. 2005. What We Know: On the Universals of Language and Rights. *Boston Review*, July 5, 2005.

Chomsky, Noam. 2006. *Language and Mind*, 3rd edition. Cambridge: Cambridge University Press.

Clarke, Samuel. 1978. *The Works*, vol. 3. New York: Garland.

Clarke, Samuel. 1998. *Demonstration of the Being and Attributes of God*. ed. Ezio Vailiati. Cambridge: Cambridge University Press.

Clemenson, David. 2005. Descartes' Direct Realisms. Unpublished manuscript, presented at the APA Pacific Division meeting, 2005.

Cook, Monte. 1987. Descartes' Alleged Representationalism. *History of Philosophy Quarterly* 4 (2): 179–195.

Cook, Monte. 2002. Robert Desgabets's Representation Principle. *Journal of the History of Philosophy* 40 (2): 189–200.

Cook, Monte. 2008. Desgabets as a Cartesian Empiricist. *Journal of the History of Philosophy* 46 (4): 501–516.

Costa, Michael. 1983. What Cartesian Ideas Are Not. *Journal of the History of Philosophy* 21 (4): 537–549.

Cottingham, John. 1986. *Descartes*. New York: Blackwell.

Cottingham, John, ed. 1994. *Reason, Will, and Sensation: Studies in Descartes' Metaphysics*. Oxford: Clarendon Press.

Cousin, Victor. 1945. Procès-verbal d'une séance d'une société cartésienne qui s'était formée à Paris dans la seconde moitié du XVIIe siècle. In *Fragments de Philosophie Cartésienne*. Paris: Charpentier.

Cronin, Timothy. 1966. *Objective Being in Descartes and in Suarez*. Rome: Gregorian University Press.

Cummins, Phillip and Guenter Zoeller, eds. 1992. *Minds, Ideas, and Objects: Essays on the Theory of Representation in Modern Philosophy*. Atascadero, CA: Ridgeview.

Curley, Edwin and Pierre-François, Moreau. 1990. *Spinoza: Issues and Directions. The Proceedings of the Chicago Spinoza Conference (1986)*. Leiden: E.J. Brill, *Brill's Studies in Intellectual History* 14: 124–135.

Daniel, Stephen. 2000. Berkeley, Suarez, and the 'Esse-Existere' Distinction. *American Catholic Philosophical Quarterly* 74 (4): 621–636.

Daston, Lorraine and Katherine Park. 1998. *Wonders and the Order of Nature 1150–1750*. New York: Zone Books.

Davidson, Donald. 1963. Actions, Reasons, and Causes. *Journal of Philosophy* 60 (3): 685–700.

Davidson, Donald. 1976. Hume's Cognitive Account of Pride. *Journal of Philosophy* 73 (4): 744–757.

Della Porta, Giambattista. 1669. *Natural Magick*. London: John Wright.

Della Rocca, Michael. 1996a. *Representation and the Mind–Body Problem in Spinoza*. Oxford: Oxford University Press.

Della Rocca, Michael. 1996b. Spinoza's Metaphysical Psychology. In *The Cambridge Companion to Spinoza*, ed. Don Garrett, 192–266. Cambridge: Cambridge University Press.

Dennett, Daniel. 1988. Quining Qualia. In *Consciousness in Contemporary Science*, eds. A. Marcel and E. Bisiach. Oxford: Oxford University Press.

Dennett, Daniel. 1991. *Consciousness Explained*. Boston, MA: Little, Brown & Company.

Dennett, Daniel. 2005. *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*. Cambridge, MA: MIT Press.

Dennett, Daniel, ed. 1987. *The Philosophical Lexicon*, 8th edition. Newark, DE: American Philosophical Association.

Descartes, René. 1649/1989. *The Passions of the Soul,* trans. Stephen Voss. Indianapolis, IN: Hackett Publishing Company.

Descartes, René. 1911. *The Philosophical Works of Descartes*, vol. I, transl. Elizabeth S. Haldane and G.R.T. Ross. Cambridge: Cambridge University Press.

Descartes, René. 1984/1985/1991. *The Philosophical Writings of Descartes*, vol. I, II, transl. J. Cottingham, R. Stoothoff, and D. Murdoch, and vol. III, transl. J. Cottingham, R. Stoothoff, D. Murdoch, and A. Kenny. Cambridge: Cambridge University Press.

Descartes, René. 1996. *Oeuvres de Descartes*, eds. Charles Adam and Paul Tannery, 11 vols. Paris: J.Vrin.

Desgabets, Robert. 1675. *Critique de la Recherche de la Vérité*. Paris.

Desgabets, Robert. 1675/1983. *Oeuvres philosophiques inédites*, Analecta Cartesiana 2, ed. J. Beaude, intro. G. Rodis-Lewis. Amsterdam, the Netherlands: Quadratures.

Desgabets, Robert. 1675/1983. *Supplément à la philosophie de Monsieur Descartes*. In *Oeuvres philosophiques inédites*, Analecta Cartesiana 2, ed. J. Beaude. Amsterdam, the Netherlands: Quadratures.

Eamon, William. 1993. *Science and the Secrets of Nature: Books of Secrets in Medieval and Early Modern Culture*. Princeton, NJ: Princeton University Press.

Easton, Patricia. 2001/2006. Robert Desgabets. In *The Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/desgabets.

Easton, Patricia. 2002. Robert Desgabets. In *The Companion to Early Modern Philosophy*, ed. Steven Nadler. Oxford: Blackwell.

Elungu, P.E. 1973. *Etendue et Connaissance dans la Philosophie de Malebranche*. Paris: Vrin.

Findlen, Paula, ed. 2004. *Athanasius Kircher: The Last Man Who Knew Everything.* New York: Routledge.

Fogelin, Robert. 1985. *Hume's Skepticism in the Treatise of Human Nature*. London: Routledge & Kegan Paul.

Foot, Philippa. 1978. *Virtues and Vices and Other Essays in Moral Philosophy*. Berkeley, CA/Los Angeles, CA: University of California Press.

Frankfurt, Harry. 1978. Descartes on the Consistency of Reason. In *Descartes: Critical and Interpretive Essays*, ed. M. Hooker. Baltimore, MD: Johns Hopkins University Press.

Gallois, André. 1998. *Occasions of Identity*. Oxford: Oxford University Press.

Garber, Daniel and Steven Nadler, eds. 2005. *Oxford Studies in Early Modern Philosophy*, vol. 2. New York: Oxford University Press.

Gardner, Howard. 1985. *The Mind's New Science.* New York: Basic Books.

Garrett, Don. 1981. Hume's Self-Doubts About Personal Identity. *The Philosophical Review* 90 (3): 337–358.

Garrett, Don. 1994. Spinoza's Theory of Metaphysical Individuation. In *Individuation in Early Modern Philosophy*, eds. Jorge Gracia and Kenneth Barber, 73–101. Albany, NY: State University of New York Press.

Garrett, Don. 2002. Spinoza's *Conatus* Argument. In *Spinoza: Metaphysical Themes*, eds. John I. Biro and Olli Koistinen, 127–158. Oxford: Oxford University Press.

Garrett, Don. 2003. Locke on Personal Identity, Consciousness, and 'Fatal Errors'. *Philosophical Topics* 31 (1–2): 95–125.

Garrett, Don. 2006. Hume's Naturalistic Theory of Representation. *Synthese* 152 (3): 301–319.

Garrett, Don. 2008. The Essence of the Body and the Part of the Mind That Is Eternal. In *A Companion to Spinoza's Ethics*, ed. Olli Koistinen. Cambridge: Cambridge University Press.

Garrett, Don, ed. 1996. *The Cambridge Companion to Spinoza*. Cambridge: Cambridge University Press.

Gaukroger, Stephen. 1995. *Descartes: An Intellectual Biography*. Oxford: Clarendon Press.

Gaukroger, Stephen, ed. 2000. *Descartes' Natural Philosophy*. London: Routledge.

Gaukroger, Stephen. 2006. *The Blackwell Guide to Descartes' Meditations*. Malden, MA: Blackwell.

Geach, P.T. 1972. *Logic Matters*. Berkeley, CA/ Los Angeles, CA: University of California Press.

Geach, P.T. 1973. Ontological Relativity and Relative Identity. In *Logic and Ontology*, ed. M.K. Munitz. New York: New York University Press.

Gennaro, Rocco J. and Charles H. Huenemann, eds. 1999. *New Essays on the Rationalists*. New York and Oxford: Oxford University Press.

Green, Michael. 1999. The Idea of a Momentary Self and Hume's Theory of Personal Identity. *British Journal for the History of Philosophy* 7 (1): 103–122.

Greenblatt, Stephen. 1991. *Marvelous Possessions: The Wonder of the New World*. Chicago, IL: University of Chicago Press.

Grice, Paul. 1991. *The Conception of Value*, intro. Judith Baker. Oxford: Clarendon Press.

Guéroult, Martial. 1952/1984. *Descartes' Philosophy Interpreted According to the Order of Reasons: The Soul and God*, transl. Roger Ariew. Minneapolis, MN: University of Minnesota Press.

Guéroult, Martial. 1963/1984–1985. *Descartes's Philosophy Interpreted According to the Order of Reasons I&II*: *The Soul and God/The Soul and Body*. Transl. Roger Ariew. Minneapolis, MN: University of Minnesota Press.

Guéroult, Martial. 1987. *Étendue et Psychologie chez Malebranche*. Paris: Vrin.

Hacking, Ian. 1975. *Why Does Language Matter to Philosophy?* Cambridge: Cambridge University Press.

Harman, Gilbert. 1990 The Intrinsic Quality of Experience. *Philosophical Perspectives* 4: 31–52.

Hatfield, Gary. 1990. *The Natural and the Normative: Theories of Spatial Perception from Kant to Helmholtz*. Cambridge, MA: MIT Press/Bradford Press.

Heinämaa, Sara, Vili Lähteenmäki, and Pauliina Remes, eds. 2007. *Consciousness: From Perception to Reflection in the History of Philosophy*. Dordrecht, the Netherlands: Springer Verlag.

Hight, Marc and Walter Ott. 2004. The New Berkeley. *Canadian Journal of Philosophy* 34 (1): 1–24.

Hoffman, Paul. 1996. Descartes on Misrepresentation. *Journal of the History of Philosophy* 34 (3): 357–381.

Hoffman, Paul. 2002. Direct Realism, Intentionality, and the Objective Being of Ideas. *Pacific Philosophical Quarterly* 83 (1): 163–179.

Holden, Thomas. 2004. *The Architecture of Matter*. Oxford: Clarendon Press.

Hooker, Michael, ed. 1978. *Descartes: Critical and Interpretive Essays*. Baltimore, MD: Johns Hopkins University Press.

Hume, David. 1985. *David Hume: Essays Moral, Political, and Literary*, ed. Eugene Miller. Indianapolis, IN: Liberty Classics.

Hume, David. 1999. *An Enquiry concerning Human Understanding*, ed. Tom L. Beauchamp. Oxford: Oxford University Press.

Hume, David. 2000. *A Treatise of Human Nature*, eds. David Fate Norton and Mary J. Norton. Oxford: Oxford University Press.

Husserl, Edmund. 1970. *The Crisis of the European Sciences and Transcendental Phenomenology*, transl. D. Carr. Evanston, IL: Northwestern University Press.

Inoue, Haruoko. 2003. The Origin of the Indirect Passions in the *Treatise*: An Analogy Between Books 1 and 2. *Hume Studies* 29: 205–221.

Jackson, Frank. 2004. Representation and Experience. In *Representation in Mind: New Approaches to Mental Representation*, eds. H. Clapin, P. Staines, and P. Slezak, 107–124. Amsterdam, the Netherlands: Elsevier.

James, Susan. 1997. *Passion and Action: The Emotions in 17th Century Philosophy*. Oxford: Clarendon Press.

Jarrett, Charles. 1990. The Development of Spinoza's Conception of Immortality. In *Dio, l'uomo, la liberta: Studi sul Breve Trattato di Spinoza*, Actes du colloque de l'Aquila, 20–23 October, 1987, ed. Filippo Mignini, 147–188. L'Aquila–Roma: Japadre.

Jarrett, Charles. 1999. Teleology and Spinoza's Doctrine of Final Causes. In *Desire and Affect: Spinoza as Psychologist*. Papers Presented at the Third Jerusalem Conference (*Ethica* III), ed. Yirmiyahu Yovel, 3–23. New York: Little Room Press.

Jarrett, Charles. 2002. Spinoza on the Relativity of Good and Evil. In *Spinoza: Metaphysical Themes*, eds. Olli Koistinen and John Biro, 159–181. Oxford: Oxford University Press.

Jolley, Nicholas. 1990. *The Light of the Soul*. Oxford: Clarendon Press.

Jolley, Nicholas. 1995. Sensation, Intentionality, and Animal Consciousness: Malebranche's Theory of the Mind. *Ratio* 7: 128–142.

Jolley, Nicholas. 1997. Introduction. *Entretiens sur la Métaphysique et sur la Religion*. In *Oeuvres Complètes de Malebranche*, 4th edition, vol. 12, ed. A. Robinet, viii–xxxiv. Paris: J. Vrin, 1991.

Jolley, Nicholas. 2000. Malebranche on the Soul. In *The Cambridge Companion to Malebranche*, ed. Steven Nadler, 31–58. Cambridge: Cambridge University Press.

Kant, Immanuel. 1984. *Critique of Pure Reason*, intro. A. D. Lindsay; transl. J. M. D. Meiklejohn. London: Dent.

Kant, Immanuel. 1993. *Grounding for the Metaphysics of Morals*, 3rd edition, transl. James W. Ellington. Indianapolis, IN: Hackett.

Kant, Immanuel. 1997. *Lectures on Metaphysics*, transl. Karl Ameriks. Cambridge: Cambridge University Press.

Kenny, Anthony. 1968. *Descartes: A Study of His Philosophy*. Oxford: Oxford University Press.

Kenny, Anthony. 1986. *Empiricism and Idealism*. Oxford: Oxford University Press.

Kirwan, Christopher. 1903. Un Ram Eau Oublié du Cartésianism. *Revue Thomiste*.

Koistinen, Olli and John Biro, eds. 2002. *Spinoza: Metaphysical Themes*. Oxford: Oxford University Press.

Kremer, Elmar, ed. 1996. *Interpreting Arnauld*. Toronto, ON: University of Toronto Press.

Leibniz, Gottfried W. 1978. *Die Philosophischen Schriften von Gottfried Wilhelm Leibniz*, ed. C.I. Gerhardt, 7 vols. Hildesheim, Germany: Georg Olms.

Leibniz, Gottfried W. 1989. *Philosophical Essays*, eds. Roger Ariew and Daniel Garber. Indianapolis, IN: Hackett.

Lennon, Thomas. 1974. The Inherence Pattern and Descartes' *Ideas*. *Journal of the History of Philosophy* 12 (1): 43–52.

Lennon, Thomas. 1980. *A Philosophical Commentary*. In Malebranche, Nicolas. *The Search After Truth*, 755–848. Columbus, OH: The Ohio State University Press.

Lennon, Thomas. 1992. Malebranche's Argument for Ideas and Its Systematic Importance. In *Minds, Ideas, and Objects: Essays on the Theory of Representation in Modern Philosophy*, Volume 2 of the *North American Kant Society Studies in Philosophy*, 57–71. Atascadero, CA: Ridgeview.

Lennon, Thomas. 1993. *The Battle of Gods and Giants: The Legacies of Descartes and Gassendi, 1615–1755*. Princeton, NJ: Princeton University Press.

Lennon, Thomas. 1998. The Cartesian Dialectic of Creation. In *The Cambridge History of Seventeenth Century Philosophy*, vol. 2, eds. Daniel Garber and Michael Ayers, 331–362. Cambridge: Cambridge University Press.

Lennon, Thomas and Patricia Easton. 1992. *The Cartesian Empiricism of François Bayle*. New York: Garland.

Lennon, Thomas and Robert Stainton, eds. 2007. *The Achilles of Rational Psychology*. Dordrecht, the Netherlands: Springer Verlag.

Locke, John. 1690/1959. *An Essay Concerning Human Understanding.* Annotated by A.C. Fraser. New York: Dover Publications.

Locke, John. 1823. *The Works of John Locke*, 10 vols. London: Thomas Tegg.

Locke, John. 1975. *An Essay Concerning Human Understanding*, ed. Peter H. Nidditch. Oxford: Clarendon Press.

Lolordo, Antonia. 2005. Descartes and Malebranche on Thought, Sensation, and the Nature of the Mind. *Journal of the History of Philosophy* 43 (4): 387–402.

Luce, A.A. 1970. Another look at Berkeley's Notebooks. *Hermathena* 110: 5–23.

Lycan, William. 1988. In Defense of the Representational Theory of Qualia (Replies to Neander, Rey, and Tye). *Philosophical Perspectives: Language, Mind, and Ontology* 12: 480–487.

Lycan, William. 2000. *Philosophy of Language: A Contemporary Introduction.* New York: Routledge.

Lycan, William. 2004. Representational Theories of Consciousness. *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/consciousness-representational/Accessed 24 December 2004.

MacIntyre, Jane. 2000. Hume's Passions: Direct and Indirect. *Hume Studies* 26: 77–86.

MacKenzie, Ann W. 1989. Descartes on Life and Sense. *Canadian Journal of Philosophy* 19 (2): 163–192.

MacKenzie, Ann W. 1990. Descartes on Sensory Representation: A Study of the *Dioptrics*. *Canadian Journal of Philosophy* 16: 109–147.

Mackie, J.L. 1977. *Ethics. Inventing Right and Wrong*. London: Penguin Books.
Maier, Anneliese. 1982. The Theory of the Elements and the Problem of Their Participation in Compounds. In *On the Threshold of Exact Science*, transl./ ed. Seven Sargent, 124–142. Philadelphia, PA: University of Pennsylvania Press.
Malebranche, Nicolas. 1674. *The Search After Truth*, ed. T. Lennon and P. Olscamp. Cambridge University Press 1997.
Malebranche, Nicolas. 1958/1967/1994. *Oeuvres Complètes de Malebranche*, ed. A. Robinet, 20 vols. Paris: J. Vrin.
Malebranche, Nicolas. 1977. *Dialogues on Metaphysics and on Religion*, transl./ eds. Nicholas Jolley and David Scott. New York: Cambridge University Press.
Malebranche, Nicolas. 1980/1997. *The Search after Truth*, transl./ed., Thomas Lennon and Paul Olscamp. Columbus, OH: The Ohio State University Press; Cambridge: Cambridge University Press.
Martin, Raymond and John, Barresi. 2002. *The Naturalization of the Soul: Self and Personal Identity in the Eighteenth Century*. London: Routledge.
Maull, Nancy. 1980. Cartesian Optics and the Geometrization of Nature. In *Descartes: Philosophy, Mathematics and Physics*, ed. Stephen Gaukroger, 23–40. Sussex: The Harvester Press.
McRae, Robert. 1972. Descartes' Definition of Thought. In *Cartesian Studies*, ed. R. J. Butler, 55–70. Oxford: Blackwell.
McRae, aRobert. 1975. On Being Present to the Mind: A Reply. *Dialogue* 14 (4): 664–666.
Mignini, Filippo, ed. 1990. *Dio, l'uomo, la liberta: Studi sul Breve Trattato di Spinoza*, Actes du colloque de l'Aquila, 20–23 October 1987. L'Aquila-Roma: Japadre.
Mijuskovic, Ben L. 1974. *The Achilles of Rationalist Arguments. The Simplicity, Unity, and Identity of Thought and Soul from the Cambridge Platonists to Kant: A Study in the History of an Argument*. The Hague: Martinus Nijhof.
Miller, Jon. 2005. Spinoza's Axiology. *Oxford Studies in Early Modern Philosophy*, 2: 149–172.
Morrisroe, Michael, Jr. 1973. Did Hume Read Berkeley? A Conclusive Answer. *Philological Quarterly* 52 (2): 310–315.
Moyal, Georges, ed. 1991. *René Descartes: Critical Assessments*, 4 vols. London: Routledge.
Muehlmann, Robert. 1992. *Berkeley's Ontology*. Indianapolis, IN: Hackett.
Muehlmann, Robert. 1995. The Substance of Berkeley's Philosophy. In *Berkeley's Metaphysics: Structural, Interpretive, and Critical Essays*, ed. Robert Muehlmann, 89–105. University Park, PA: Pennsylvania State University Press.
Myro, George. 1985. Identity and Time. In *Grounds of Rationality: Intentions Categories and Ends*, eds. Richard Grandy and Richard Warner, 383–409. Oxford: Oxford University Press.
Myro, George. 1986. Time and Essence. *Midwest Studies in Philosophy* 11: 331–341.
Nadler, Steven. 1989. *Arnauld and the Cartesian Philosophy of Ideas*. Princeton, NJ: Princeton University Press.

Nadler, Steven. 1992a. *Malebranche and Ideas*. Oxford: Oxford University Press.
Nadler, Steven. 1992b. Intentionality in the Arnauld–Malebranche Debate. In *Minds, Ideas, and Objects: Essays on the Theory of Representation in Modern Philosophy*, Volume 2 of the *North American Kant Society Studies in Philosophy*, 73–84. Atascadero, CA: Ridgeview.
Nadler, Steven. 1994a. Malebranche's Theory of Perception. In *The Great Arnauld and Some of His Philosophical Correspondents*, ed. Elmar Kremer, 108–128. Toronto, ON: University of Toronto Press.
Nadler, Steven. 1994b. Descartes and Occasional Causation. *British Journal for the History of Philosophy* 2: 35–54.
Nadler, Steven. 2001. *Spinoza's Heresy*. Oxford: Oxford University Press.
Nelson, Alan. 1996. The Falsity in Sensory Ideas: Descartes and Arnauld. In *Interpreting Arnauld*, ed. Elmar Kremer. Toronto, ON: University of Toronto Press.
Normore, Calvin. 1986. Meaning and Objective Being: Descartes and His Sources. In *Essays on Descartes' Meditations*, ed. Amèlie Oksenberg Rorty, 223–241. Berkeley, CA: University of California Press.
Norton, David F., ed. 1993. *The Cambridge Companion to Hume*. Cambridge: Cambridge University Press.
O'Connor, Timothy and Hong Yu W. 2006. Emergent Properties. Entry in *The Stanford Encyclopedia of Philosophy (Winter 2006 Edition)*, ed. Edward N. Zalta. http://plato.stanford.edu/archives/win2006/entries/properties-emergent/
O'Neil, Brian. 1974. *Epistemological Direct Realism in Descartes' Philosophy*. Albuquerque, NM: University of New Mexico Press.
O'Neill, Eileen. 1987. Mind–Body Interaction and Metaphysical Consistency: A Defense of Descartes. *Journal of the History of Philosophy* 25 (1): 227–245.
Pappas, George. 1992. Perception of the Self. *Hume Studies* 18 (2): 275–280.
Pasnau, Robert. 1997. *Theories of Cognition in the Later Middle Ages*. Cambridge: Cambridge University Press.
Penelhum, Terence. 1955. Hume on Personal Identity. *Philosophical Review*. 64: 571–589.
Penelhum, Terence. 1993. Hume's Moral Psychology. In *Cambridge Companion to Hume*, ed. David Fate Norton, 117–147. Cambridge: Cambridge University Press.
Pessin, Andrew. 2003. Descartes' Nomic Concurrentism: Finite Causation and Divine Concurrence. *Journal of the History of Philosophy* 41 (1): 25–49.
Pessin, Andrew. 2004. Malebranche on Ideas. *Canadian Journal of Philosophy* 34 (2): 241–286.
Pessin, Andrew. 2007. Descartes's Theory of Ideas. *Stanford Encyclopedia of Philosophy*.
Pitcher, George. 1977. *Berkeley*. London: Routledge & Kegan Paul.
Pitson, A.E. 2002. *Hume's Philosophy of the Self*. London: Routledge.
Popkin, Richard H. 1959. Did Hume Ever Read Berkeley? *Journal of Philosophy* 56: 992–995.

Popkin, Richard H. 1964. So, Hume Did Read Berkeley. *Journal of Philosophy* 61: 773–779.
Purviance, Susan M. 1997. The Moral Self and the Indirect Passions. *Hume Studies* 23: 195–212.
Putnam, Hilary. 1981. *Reason, Truth and History*. Cambridge: Cambridge University Press.
Pyle, Andrew. 2003. *Malebranche*. New York: Routledge.
Radner, Daisie. 1978. *Malebranche: A Study of a Cartesian System*. Assen, the Netherlands: Van Gorcum.
Radner, Daisie. 1988. Thought and Consciousness in Descartes. *Journal of the History of Philosophy* 36 (3): 439–452.
Radner, Daisie. 1994. Malebranche and the Individuation of Perceptual Objects. In *Individuation and Identity in Early Modern Philosophy*, eds. Kenneth Barber and Jorge Gracia, 59–72. Albany, NY: SUNY Press.
Raynor, David R. 1980. '*Minima Sensibilia*' in Berkeley and Hume. *Dialogue* 19 (1): 196–199.
Raynor, David R. 1990. Hume and Berkeley's *Three Dialogues*. In *Studies in the Philosophy of the Scottish Enlightenment*, ed. M. A. Stewart. Oxford: Oxford University Press.
Reid, Jasper. 2003. Malebranche on Intelligible Extension. *British Journal for the History of Philosophy* 11 (4): 581–608.
Ricci, Mateo. 1610/1953. *China in the Sixteenth Century: The Journals of Matthew Ricci 1583–1610*, transl. Louis Gallagher. New York: Random House.
Robinet, André. 1957. *Correspondence Leibniz-Clarke; présentée d'après les manuscrits originaux des bibliothèques de Hanovre et de Londres*. Paris: Presses Universitaires de France.
Rodis-Lewis, Geneviève. 1993. Der Cartesianismus in Frankreich. In *Die Philosophie Des 17. Jarhunderts, Band 2, Frankreich und Niederlande*, ed. Jean-Pierre Schobinger, 423–431. Basel: Schwabe.
Rome, Beatrice. 1963. *The Philosophy of Malebranche: A Study of His Integration of Faith, Reason, and Experimental Observation*. Chicago, IL: Henry Regnery.
Rorty, Amèlie Oksenberg. 1982. From Passions to Emotions and Sentiments. *Philosophy* 57: 159–72.
Rorty, Amèlie Oksenberg, ed. 1986. *Essays on Descartes' Meditations*. Berkeley, CA: University of California Press.
Rorty, Richard. 1979. *Philosophy and the Mirror of Nature*. Princeton, NJ: Princeton University Press.
Rosenberg, Alexander. 1993. Hume and the Philosophy of Science. In *The Cambridge Companion to Hume*, ed. David Fate Norton, 64–89. Cambridge: Cambridge University Press.
Rozemond, Marleen. 1998. *Descartes' Dualism*. Cambridge, MA: Harvard University Press.
Rozemond, Marleen. 1999. Descartes on Mind–Body Interaction: What's the Problem? *Journal of the History of Philosophy* 37 (3): 435–467.

Rozemond, Marleen. 2006. The Nature of the Mind. In *The Blackwell Guide to Descartes' Meditations*, ed. Stephen Gaukroger, 48–66. Malden, MA: Blackwell.

Rozemond, Marleen. 2007. The Achilles Argument and the Nature of Matter in the Clarke–Collins Correspondence. In *The Achilles of Rational Psychology*, eds. Tom Lennon and Robert Stainton. Dordrecht, the Netherlands: Springer.

Rozemond, Marleen and Gideon, Yaffe. 2004. Peach Trees, Gravity and God: Locke on Mechanism. *British Journal for the History of Philosophy* 12: 387–412.

Schmaltz, Tad. 1992. Sensation, Occasionalism, and Descartes' Causal Principles. In. *Minds, Ideas, and Objects: Essays on the Theory of Representation in Modern Philosophy*, eds. Phillip Cummins and Guenter Zoeller, 37–56. Atascadero, CA: Ridgeview.

Schmaltz, Tad. 1995. Malebranche's Cartesianism and Lockean Colours. *History of Philosophy Quarterly* 12: 387–403.

Schmaltz, Tad. 1996. *Malebranche's Theory of the Soul: A Cartesian Interpretation*. New York: Oxford University Press.

Schmaltz, Tad. 1999. Causation and Similarity in Descartes. In *New Essays on the Rationalists*, eds. Rocco J. Gennaro and Charles Huenemann. Oxford: Oxford University Press.

Schmaltz, Tad. 2002. *Radical Cartesianism: Developments in Desgabets and Regis*. Cambridge: Cambridge University Press.

Schmitter, Amy M. 2000. The Wax and I: Perceptibility and Modality in the Second Meditation. *Archiv für Geschichte der Philosophie* 82 (1): 178–201.

Seager, William. 1999. *Theories of Consciousness*. New York: Routledge.

Shoemaker, Sidney. 1986. Introspection and the Self. In *Studies in the Philosophy of Mind*, eds. Peter A. French, Theodoure E. Vehlin, and Howard K. Wettstein, 101–120. Minneapolis, MN: Midwest Studies in Philosophy 10.

Shumaker, Wayne. 1989. *Natural Magic and Modern Science: Four Treatises, 1590–1657*. Binghamton, NY: Center for Medieval and Early Renaissance Study, SUNY.

Siewert, Charles, "Consciousness and Intentionality," *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, ed. Edward N. Zalta. http://plato.stanford.edu/archives/fall2008/entries/consciousness-intentionality/

Simmons, Alison. 1999. Are Cartesian Sensations Representational? *Nous* 33 (3): 347–369.

Simmons, Alison. 2001. Sensible Ends: Latent Teleology in Descartes' Account of Sensation. *Journal of the History of Philosophy* 39 (1): 49–75.

Simmons, Alison. 2003. Descartes on the Cognitive Structure of Sensory Experience. *Philosophy and Phenomenological Research* 67 (3): 549–579.

Spinoza, Benedictus de. 1972. *Spinoza Opera*, ed. Carl Gebhardt, 4 vols. Reprint from 1925. Heidelberg: Carl Winter.

Spinoza, Benedictus de. 1985. *The Collected Works of Spinoza*, *Volume I*, transl./ed. Edwin M. Curley. Princeton, NJ: Princeton University Press.

Spinoza, Benedictus de. 2001. *Theological–Political Treatise*, transl. Seymour Feldman, ed. Samuel Shirley. Indianapolis, IN: Hackett.

Spinoza, Benedictus de. 2002. *Spinoza: Complete Works*, transl. Samuel Shirley, ed./intro. Michael L. Morgan. Indianapolis, IN/Cambridge: Hackett.

Stafford, Barbara. 1994. *Artful Science: Enlightenment Entertainment and the Eclipse of Visual Education*. Cambridge, MA: MIT Press.

Stevenson, Gordon Park. 1998. Humean Self-Consciousness Explained. *Hume Studies* 24 (1): 95–129.

Stewart, M. A., ed. 1990. *Studies in the Philosophy of the Scottish Enlightenment*. Oxford: Oxford University Press.

Stroud, Barry. 1977. *Hume*. London: Routledge.

Swift, Jonathan. 1715. *Gulliver Travels* in *The Writings of Jonathan Swift*, ed. R. Greenberg and W. Piper. W.H. Norton and Company, New York 1973.

Swift, Jonathan. 1973. *The Writings of Jonathan Swift*, ed. Robert Greenberg. New York: Norton Critical Editions.

Terpak, Francis. 2001. *Devices of Wonder: From the World in a Box to Images on a Screen*. Los Angeles, CA: Getty Publications.

Thiel, Udo. 1994. Hume's Notions of Consciousness and Reflection in Context. *British Journal for the History of Philosophy* 2 (2): 75–115.

Tipton, Ian. 1966. Berkeley's View of Spirit. *New Studies in Berkeley's Philosophy*, ed. Warren E. Steinkraus. New York: Hold, Rinehard, and Winston.

Tipton, Ian. 1974. *Berkeley: The Philosophy of Immaterialism*. London: Methuen.

Tipton, Ian. 1992. 'Ideas' and 'Objects': Locke on Perceiving 'Things'. In. *Minds, Ideas, and Objects: Essays on the Theory of Representation in Modern Philosophy*, eds. Phillip Cummins and Guenter Zoeller, 97–110. Atascadero, CA: Ridgeview.

Traiger, Saul. 1987. Impressions, Ideas and Fictions. *Hume Studies* 13: 381–399.

Turbayne, Colin M. 1959. Berkeley's Two Concepts of Mind. *Philosophy and Phenomenological Research* 20 (1): 85–92.

Turbayne, Colin M. 1962. Berkeley's Two Concepts of Mind, Part II. *Philosophy and Phenomenology Research* 22 (3): 383–386.

Tye, Michael. 2000. *Consciousness, Color, and Content*. Cambridge, MA: MIT Press.

Unger, Peter. 2006. *All the Power in the World*. Oxford: Oxford University Press.

Vailiati, Ezio. 1993. Clarke's Extended Soul. *Journal of the History of Philosophy* 31 (2): 387–403.

Vailiati, Ezio. 1997. *Leibniz and Clarke: A Study of Their Correspondence*. Oxford: Oxford University Press, 1997.

Van Fraassen, Bas C. 2004. Armstrong on Laws and Probabilities. In *Readings on Laws of Nature*, ed. John W. Carroll. Pittsburgh, PA: University of Pittsburgh Press.

Viljanen, Valtteri. 2007. *Spinoza's Dynamics of Being: The Concept of Power and Its Role in Spinoza's Metaphysics*. Reports from the Department of Philosophy, vol. 19. Turku: University of Turku.

Vinci, Thomas. 1998. *Cartesian Truth*. Oxford: Oxford University Press.

Watson, Richard. 1987. *The Breakdown of Cartesian Metaphysics*. Atlantic Highlands, NJ: Humanities Press International.

Waxman, Wayne. 1994. *Hume's Theory of Consciousness*. Cambridge: Cambridge University Press.
Wells, Norman. 1967. Objective Being: Descartes and His Sources. *The Modern Schoolman* 45: 49–61.
Wells, Norman. 1990. Objective Reality of Ideas in Descartes, Caterus, and Suárez. *Journal of the History of Philosophy* 28 (1): 33–61.
Wiggins, David. 2001. *Sameness and Substance Renewed*. Cambridge: Cambridge University Press.
Williston, Byron and André Gombay, eds. 2003. *Passion and Virtue in Descartes*. New York: Humanity Books/Prometheus Books.
Wilson, Margaret. 1978. *Descartes*. London: Routledge.
Wilson, Margaret. 1979. Superadded Properties: The Limits of Mechanism in Locke. *American Philosophical Quarterly* 16: 143–150.
Wilson, Margaret. 1982. Superadded Properties: A Reply to M.R. Ayer. *The Philosophical Review* 91 (2): 247–252.
Wilson, Margaret. 1990. Descartes on the Representationality of Sensation. Reprinted in *Ideas and Mechanism: Essays on Early Modern Philosophy*, ed. Margaret Wilson, 69–83. Princeton, NJ: Princeton University Press.
Wilson, Margaret. 1991. Descartes on the Origin of Sensation. *Philosophical Topics* 19: 293–323.
Wilson, Margaret. 1992. Descartes on the Perception of Primary Qualities. Reprinted in *Ideas and Mechanism: Essays on Early Modern Philosophy*, ed. Margaret Wilson, 26–40. Princeton, NJ: Princeton University Press.
Wilson, Margaret. 1994. Descartes on Sense and 'Resemblance'. Reprinted in *Ideas and Mechanism: Essays on Early Modern Philosophy*, ed. Margaret Wilson, 10–25. Princeton, NJ: Princeton University Press.
Wilson, Margaret. 1999. *Ideas and Mechanism: Essays on Early Modern Philosophy*. Princeton, NJ: Princeton University Press.
Yolton, John. 1975. On Being Present to the Mind: A Sketch for the History of an Idea. *Dialogue* 14: 373–388.
Yolton, John. 1983. *Thinking Matter: Materialism in Eighteenth-Century Britain*. Minneapolis, MN: University of Minnesota Press.
Yolton, John. 1984. *Perceptual Acquaintance from Descartes to Reid*. Minneapolis, MN: University of Minnesota Press.
Yolton, John. 1991. *Locke and French Materialism*. Oxford: Oxford University Press.
Yovel, Yirmiyahu, ed. 1991. *God and Nature: Spinoza's Metaphysics*. Papers Presented at the First Jerusalem Conference (*Ethica* I). Leiden: E.J. Brill.
Yovel, Yirmiyahu. 1999. *Desire and Affect: Spinoza as Psychologist*. Papers Presented at the Third Jerusalem Conference (*Ethica* III). New York: Little Room Press.

Author Index*

* Many thanks to James Stuckey for compiling the indices.

Subjects Index