

Studies in History and Philosophy of Science 29

Martin Lenz
Anik Waldow *Editors*

Contemporary Perspectives on Early Modern Philosophy

Nature and Norms in Thought

 Springer

Contemporary Perspectives on Early Modern Philosophy

STUDIES IN HISTORY
AND PHILOSOPHY OF SCIENCE

VOLUME 29

General Editor:

STEPHEN GAUKROGER, *University of Sydney*

Editorial Advisory Board:

RACHEL ANKENY, *University of Adelaide*
PETER ANSTEY, *University of Otago*
STEVEN FRENCH, *University of Leeds*
KOEN VERMEIR, *Katholieke Universiteit, Leuven*
OFER GAL, *University of Sydney*
JOHN SCHUSTER, *University of New South Wales*
RICHARD YEO, *Griffith University*

For further volumes:
www.springer.com/series/5671

Martin Lenz • Anik Waldow
Editors

Contemporary Perspectives on Early Modern Philosophy

Nature and Norms in Thought

 Springer

Editors

Martin Lenz
University of Groningen
Groningen
The Netherlands

Anik Waldow
Department of Philosophy
University of Sydney
Sydney, NSW
Australia

ISSN 0929-6425

ISBN 978-94-007-6240-4

ISBN 978-94-007-6241-1 (eBook)

DOI 10.1007/978-94-007-6241-1

Springer Dordrecht Heidelberg New York London

Library of Congress Control Number: 2013936401

© Springer Science+Business Media Dordrecht 2013

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The opposition between the natural and normative provides us with important philosophical distinctions, for instance, between causes and reasons, explanation and understanding, facts and values or determination and freedom. However, these seemingly clear lines of demarcation often turn out to be rather complicated, as is particularly evident in recent attempts to ‘naturalize’ intentionality and knowledge. Are our intentional relations to the world determined by factors that are explicable in terms of natural properties and laws? Or are they rather regulated by socially established norms? The fact that we can ask these questions shows that the distinction between nature and normativity, whose origin is commonly traced back to the early modern period, is still confusing, even though we might have become very familiar with it.

This book is an attempt to clarify how early modern philosophers construed the influence of nature and norms on thought, and it tackles this task by invoking a decidedly contemporary perspective. The idea for this book was born at a workshop on “Nature versus Normativity?” held at Humboldt University, Berlin, in July 2011, made possible by a grant of the German Research Foundation.

We owe much to the institutions and people who supported our workshop and through this helped us collectively to bring our work to the present stage. Thus, we would like to express our gratitude to Dominik Perler for hosting this event as well as for his unflinching collaboration and support. We would also like to thank the Topoi Excellence Cluster and Humboldt University for having provided us with the conference venue. Special thanks to Sebastian Bender, Sanja Dembić, Luz Christopher Seiberth and Simone Ungerer for administrative and logistic support. The Australian Research Council funded the related research project “The Rise of Empiricism,” and in this way offered great support in establishing the research collaboration between Berlin and Sydney. The School of Philosophical and Historical Inquiry, University of Sydney, funded a second workshop on a related topic in Sydney in September 2011 and gave three of the contributors the opportunity for further discussion of their papers.

Our warmest thanks, also, to Stephen Gaukroger for accepting to publish our volume in this series, to Lucy Fleet for her support in finalising the manuscript, to Annette Pierdziwol for her invaluable help with the editing of many of the chapters, as well as to our group of contributors.

Sydney and Groningen

The Editors

Contents

1 Nature and Norms in Thought	1
Martin Lenz and Anik Waldow	
Part I Nature’s Influence on the Mind	
2 Intentionality Bifurcated: A Lesson from Early Modern Philosophy?	13
Lionel Shapiro	
3 Ideas as Thick Beliefs: Spinoza on the Normativity of Ideas	37
Martin Lenz	
4 Three Problems in Locke’s Ontology of Substance and Mode	51
Antonia LoLordo	
5 Kant on Imagination and the Natural Sources of the Conceptual	65
Johannes Haag	
6 Naturalized Epistemology and the Genealogy of Knowledge	87
Martin Kusch	
Part II Shaping the Norms of Our Intellectual and Practical Engagement with the World	
7 Sensibility and Metaphysics: Diderot, Hume, Baumgarten, and Herder	101
Stephen Gaukroger	
8 Back to the Facts – Herder on the Normative Role of Sensibility and Imagination	115
Anik Waldow	

9	Extending Nature: Rousseau on the Cultivation of Moral Sensibility	135
	Annette Pierdziwol	
10	The Piacular, or on Seeing Oneself as a Moral Cause in Adam Smith	159
	Eric Schliesser	
11	Explaining and Describing: Panpsychism and Deep Ecology	179
	Michael Hampe	
	Author Biographies	203
	Index	205

Chapter 1

Nature and Norms in Thought

Martin Lenz and Anik Waldow

The present volume joins contributions to early modern debates on nature and norms in thought with decidedly contemporary perspectives, thereby hoping to shed new light on developments in early modern philosophy as well as enrich current discussions on the relation between nature and norms. Clearly, the relation between mind and world poses perennial problems and debates. How do we explain that thoughts and other mental states have content? What makes it the case that some thought is about this rather than that thing? Do our perceptions and thoughts match the world? How do we categorize things? Do our concepts carve up nature at its joints? Is thinking a kind of action? Where does it take place? Is it embodied? What makes thoughts and sentences true or false? Do beliefs aim at truth? Do true beliefs constitute knowledge? What makes our thoughts adequate? Can our beliefs fail to reach epistemic goals? Does thought depend on interaction with other thinkers? Can other animals think too? Do we need language to think? Can we ever be sure about anything?

In the twentieth century, these and related debates about the relation between mind and world have taken on a new form as a result of the emerging influence of naturalist and normativist theories.¹ We can better understand what is involved in this, and especially the claim that nature and norms are to be conceived in opposition to one another, by looking at the question of how mental states acquire content. *Naturalists* such as Jerry Fodor typically answer this question by invoking *causal*

¹ See, for instance, De Caro and Macarthur (2010) and Putnam (2002).

M. Lenz (✉)

Department of the History of Philosophy, University of Groningen,
Oude Boteringestraat 52, 9712GL Groningen, The Netherlands
e-mail: m.lenz@rug.nl

A. Waldow

Department of Philosophy, University of Sydney, Sydney, NSW, Australia
e-mail: anik.waldow@sydney.edu.au

theories of mental content.² The basic idea here is that thoughts are about their causes: I can think about water because there is water in the world that causes me to have the concept of water. By contrast, *normativists* such as Robert Brandom see themselves as standing in the tradition of Wittgenstein's private language argument and argue that references to causal mechanisms are not sufficient as an explanation for content determination.³ The basic idea is that the contents of concepts are fixed by the *use* of these concepts, rather than causal processes. On this view, it is because our concepts play an essential role in inferential relations (to which thinkers subscribe) and because different thinkers interact and sanction one another in their use of language that it is possible that concepts acquire content. Conceived in this way, thinking about water presupposes a commitment to inferences that combine claims about attributes such as being liquid and wet.

Once we turn to the problem of error, the divide between naturalists and normativists becomes even more obvious. What does it mean to say that a concept has been misapplied or that a belief is false? Imagine that you express a false belief by calling the water in front of you "gin". What are the resources to explain this error? Is this error merely the result of a mismatch between your concept and the relevant facts, or rather a consequence of having broken a conventionally established rule?

Naturalists tend to think that this error is the result of a deviation from facts. But even if it is such a deviation that causes error, one may object that it is hard to explain how it is possible to identify which one among all other causal relations gives rise to the right kind of concept (namely, water). Normativists circumnavigate this problem by arguing that talk about misapplication, misrepresentation, correctness and incorrectness already betrays the omnipresence of socially established norms.

Generally speaking, while naturalists appeal to causal relations between mind and world in order to explain how it is possible to think and speak meaningfully about the world, normativists refer to inferential relations and social interactions between agents.

Of course, to think of naturalism and normativism as standing in opposition to one another is not specific to debates in the philosophy of mind, but also manifests itself in metaphysics, epistemology and ethics, especially when the question arises as to whether it is possible to naturalize these disciplines. However, despite this overarching trend in philosophy, there are also approaches that try to bridge what others conceive as a gap, namely, by naturalizing normativity itself. Ruth Millikan, for instance, has chosen this path. She neither explains misrepresentation in terms of causal processes nor by invoking a breach of social conventions, but instead conceptualizes error in relation to the evolutionary function of cognitive mechanisms. Put crudely, her account suggests that if the water in front of me produces gin-thoughts, the error does not lie in the breaking of a rule but in the relevant cognitive mechanism's failure to comply with its biological purpose.⁴ John McDowell's theory of concepts, Philippa

² See esp. Fodor (1987).

³ See Brandom (1994).

⁴ See Millikan (2004).

Foot's ethics or Peter Stemmer's ontology of normativity can, in some sense, each be read in a similar vein, namely as attempts to undermine the opposition between nature and normativity.⁵ We can here see that one crucial question that still needs to be answered is whether it is justified to think of the natural and normative as two mutually exclusive categories, and as such as two separate conceptual realms, or whether we have reasons to think that norms and nature can jointly figure in our explanations.

But what, then, are the natural and the normative aspects of thought? Although contemporary debates often focus on methodological and metaphilosophical issues in relation to the function of explanation (its placement and fruitfulness), much of what is going on here is a consequence of the conceptual divide between the natural and the normative – a divide that has often been seen as having its roots in early modern philosophy. Thus, the conceptual question of the compatibility of the natural and the normative clearly has a historical dimension.⁶ In Descartes, for instance, the separation of nature from the realm of the normative takes the form of an exclusion of teleological explanations from the world of mechanical causes, and also informs the ubiquitous and difficult attempts to distinguish between sensation as a mechanical process, on the one hand, and judgment as a norm-guided act of the intellect, on the other. Another infamous way of separating the natural from the normative is typically associated with the Humean distinction between *Is* and *Ought*, a distinction that has often been discussed in connection with the Moorean charge of the naturalistic fallacy.

But although these examples are present in the history of philosophy, we nevertheless must ask whether it is justified to conceive of these early modern approaches as “forerunners” of modern theories with a strong focus on the opposition between the natural (taken as the realm of law) and the normative (taken as something non-natural that mainly rests on social conventions). Richard Rorty's historiography suggests that early modern philosophy may be conceived as having introduced the rigid distinction between the realms of the natural and normative, but that it did so in a very confused and erroneous way in its failure to see that mental representations were not simply “mirroring” nature.⁷

Rorty's historiography is still fairly influential,⁸ even though it might appear to some like an “outdated caricature” of early modern philosophy of mind and epistemology.⁹ Given its influence, however, should we not worry that this book too will appear anachronistic? It might look as if the entire debate in contemporary philosophy about the standing of nature and the normative has been inspired by the pragmatist tradition and as such constitutes a very recent development in philosophy.

⁵ See McDowell (1986), Foot (2003), and Stemmer (2008).

⁶ See, for instance, Brandom (2002).

⁷ See Rorty (1979, 131–164).

⁸ See, for instance, Tartaglia (2007).

⁹ Hatfield (2005, 98).

To be sure, the danger of providing an anachronistic account emerges for anyone working in the history of philosophy. A way to handle this problem then perhaps consists in accepting that we do indeed have philosophical interests of our own and that we want to bring them to bear on our discussion of the past, and vice versa. This book, then, is an attempt to do precisely this: it will provide studies of historical accounts by joining them with decidedly contemporary perspectives.

But what then can be said about the historical dimension of the question regarding the compatibility of the natural and the normative? Generally speaking, the first thing to note is that early modern works often *conjoin* what has only later been divided, namely the treatment of epistemological, metaphysical and moral questions.

A second point concerns the notion of nature. As has already been indicated above, Descartes' and Spinoza's exclusions of teleological explanation from the realm of nature has often been seen as indicating a strict division of the natural world of mechanism, on the one hand, and the world of human intentions, conventions and norms, on the other. To put this claim into perspective, it is important to note, however, that early modern conceptions of nature and natural laws are notoriously ambiguous. While "law of nature" (*lex naturae*) is mostly understood as referring to divine or moral law, which in its fixity contrasts with variable conventions, nature (*natura*) was also taken to be that out of which second nature (*altera natura*) emerges when acquired habits establish themselves. This suggests that it would be wrong to believe that the meaning of the concept of nature is exhaustively captured by its opposition to the concept of convention.¹⁰

A third point that ought to be considered in this context is the charge of individualism so often associated with early modern philosophy. While it is certainly true that early modern philosophers had a great interest in human beings and their history, this interest did not lead them to focus on the individual mind or agent in isolations from the social sphere. This in principle already follows from the early modern approach to philosophy in which epistemology and ethics are explicitly conjoined. Here then it would appear that much of the work ostensibly focused on the individual in fact regards intersubjective practices as a means of establishing knowledge and morals.

We can here already see – and this will become even clearer once we enter into the various discussions in this book – that early modern philosophy offers an abundance of rather elaborate views on the relation between nature and normativity, that is, views that are very different from those depicted in Rorty's caricature. A good example is Locke's discussion of essences. On the face of it, there seems to be nothing particularly normative about essences: they are merely that which causes a thing to have these rather than those properties. However, in fact it turns out that each time we take a thing to belong to a certain kind, our cognitive make-up and practical needs influence how we do this. This means that if we were beings with different capacities and needs we would carve up the world quite differently. Thus, although

¹⁰On the early modern debates on the *lex naturae* see Haakonssen (1998).

there are real essences, our actual attributions of kind membership are governed by our interests and needs. Now, since such ascriptions are not only driven by certain conventions, but are also unfailingly prescriptive in that they specify what a thing ought to be like in order to belong to a certain kind, it becomes evident that in Locke kind membership cannot be understood independently of the normative constraints that render this membership possible.¹¹

More generally, we can also perceive that in early modern philosophy the attempt to substantiate a theory about the normative dimension of ostensibly natural processes is often interlinked with the endeavor of redefining the scope and limits as well as the methodology of the sciences. For instance, Locke's discussion of essences not only reveals the normative implications of our kind terms, but also intends to restructure the foundations of the various sciences and their respective methodologies. And although it here seems that there are in fact connections between rather different philosophical intentions, it would be premature to conclude from this alone that reflections on the relation between norms and nature must determine a philosopher's choice of methodology. Spinoza, for example, is clear that the normative notions of goodness and badness do not pick out kinetic or mechanistic properties of the universe, and yet he still thinks that these notions are methodologically indispensable.¹²

In order to do justice to the complex structures undergirding the conceptions of what mind is and how it relates to the world in its thoughts and scientific explanations, this book is divided into two parts. The first part discusses the way in which nature and norms can be conceived as influencing the formation of thought. The chapters of the second part will investigate how a certain conception of what the mind is, and ought to do when engaging with the deliverances of the senses, reason and the imagination, becomes the defining feature in the shaping of the norms of philosophical inquiry. Each part will contain one discussion that examines the relevant topics in a contemporary context, thus opening up a space in which continuities (and discontinuities) between early modern perspectives and those of today can be perceived.

1.1 Part I Nature's Influence on the Mind

The chapters in this part are case studies of themes prominently discussed in contemporary philosophy of mind, epistemology and ontology. They focus on questions of intentionality, the nature of beliefs, the standard of ontological classification, the relation between perceptual and conceptual representation, and the concept of knowledge. By examining how early modern philosophers conceived of nature as underwriting the formation of normative standards and the way in which norms

¹¹ See Lenz (2010, 446–454) and Pasnau (2011, 633–661).

¹² See Spinoza (1994, *Ethica* IV praefatio).

embedded in a social context can provide a framework for nature to assert itself, it will become clear that there is reason for us to rethink and refresh our current philosophical theories.

At least since Brentano, questions of intentionality have revolved around one particular problem, which is often even named “*the* problem of intentionality”: that is, the problem of what it is for a mental state to be a representation *of* something else, let us say a thing in the world. As has already been noted, the possibility that something can be misrepresented forces naturalists to account for the normative dimension of perceptual processes. Yet, as Lionel Shapiro argues, early modern debates can be used to broaden and refine our understanding of what intentionality is. According to Shapiro, Descartes and Locke, for instance, distinguish between two kinds of ofness: the first being the ordinary kind of “propositional ofness” that accounts for our capacity to form the thought “that p”. The second kind of intentionality is a “representational ofness” the recognition of which was forced upon Descartes and Locke by the need to account for the success of our proper methods of inquiry. According to Shapiro, this distinction between two kinds of intentionality maps onto two kinds of normativity exemplified by the well-known truth norm of propositional intentionality, on the one hand, and the norms governing the ideal improvement of representational intentionality, on the other.

The idea that truth might not be the only norm that governs the intentionality of beliefs is also taken up by Martin Lenz’s discussion of Spinoza. Against Robert Brandom Lenz argues that Spinoza offers a fairly thorough account of the normativity of ideas. In construing ideas as propositional attitudes, he suggests that Spinoza’s ideas are beliefs that respond to two kind of normative constraints. On the one hand, beliefs count as naturally normative in that they are grounded in our striving for self-preservation (*conatus*). On the other hand, they exhibit a kind of socially rooted normativity in that they are governed by associations reinforced by custom and convention.

Besides accounting for intentional and epistemological aspects of thought, early modern theories of ideas also serve to explain ontological classification. As is well-known, Locke’s ontological distinctions between qualities, substances, modes and relations are driven by his taxonomy of ideas. Whether something can be called a substance is decided by the fact that it can be conceived as existing independently of something else. But what precisely does it mean to conceive of something in this way? To address this question, Antonia LoLordo invokes three problems in Locke’s analysis of the relation between substance and mode. She argues that there are crucial tensions in the way Locke construes the realm of so-called natural kinds by contrasting it with the realms of mathematics and morality.

In the Kantian context, the problem of how it is possible to bridge the gap between mind and world has often been articulated by drawing attention to the various processes involved in cognition. Sensation has here been seen as a natural affair in that it requires of the cognitive subject no more than passive receptivity, while the formation of concepts has been regarded as norm-guided in that it responds to the demands of spontaneous reason. With the introduction of such conceptual divides, it becomes difficult to see, however, how it possible to form empirical concepts at

all. Johannes Haag's paper addresses this problem by examining how on Kant's account natural and normative capacities can be conceived as interacting with one another. Haag takes off from a broadly Sellarsian framework, but goes beyond it by arguing that the possibility of such interaction depends on the imagination and its role in the mediation between sensation and conception.

As has been noted above, a common strategy in contemporary debates on the relation of nature and normativity consists in naturalizing the notions of intentionality and knowledge. One of the seemingly most radical proposals comes from Hilary Kornblith who argues (against Edward Craig and others) that knowledge is a natural kind rather than the phenomenon we commonly think to pick out in our discussions on justification and epistemic norms. By refining the historical notion of knowledge defended by Craig, Martin Kusch shows that Kornblith stands much closer to Craig than his criticism of Craig suggests, thus underlining that in this particular case naturalists and normativists can in fact be perceived as allies rather than opponents.

1.2 Part II Shaping the Norms of Our Intellectual and Practical Engagement with the World

The second part of this book will examine the ways in which in the early modern era reflections on the mind's various sentimental and rational capacities were utilized to transform the concept of the human being, and in relation to this the norms of intellectual and moral practices. Reflections of this kind usually emerged in response to questions about the place of the thinking subject in nature and society, our ability to take control of the way we experience and conceptualize the world, and the normative demands that ought to guide these processes.

As we will see in many of the chapters of this second part of the book, at the heart of this development lies a change in the perception of what reason is and how it is able to manifest itself in our thoughts and actions – a change that was often motivated by a reevaluation of the role that our sensibility plays in the formation of higher rational and moral capacities. Different thinkers had different ideas about the way in which the relation between affective and rational capacities should be understood. However, a common feature of approaches inspired by Locke's sensationalist epistemology was that mental activity – even in its most abstract and life-detached variation – was no longer regarded in isolation from spontaneous sensory and sentimental reactions, but instead was conceived as the result of the mind's ability to engage with affective processes caused by one's embodied existence.

In a certain sense, this recognition of the body as a necessary prerequisite for the development of higher rational capacities led to a conception of nature as the all encompassing category that needed to be reckoned with when trying to explain what is characteristic of human intellectual, moral and social life. The interesting development in this was that nature was often represented as the realm out of which specifically human attributes, such as reason and language, emerged to the effect that, on the one hand, human beings could be regarded as integral parts of nature

while, on the other, they could also be conceptualized as active agents able to comply with the normative demands of reason and morality.

Curiously, in this debate more was at stake than the mere reevaluation of reason and its relation to affective components. The claim that the mind must develop the right kind of skills and attitudes in order to be able to engage with the world in a fruitful way was usually expressed to articulate dissatisfaction with the current philosophical practices. Questions about the way in which agents, inquirers and minds ought to proceed when exploring the world and their own responsibilities thus translated into concerns about the aims and goals of the philosophical discipline, and called for a justification of current intellectual practices.

Stephen Gaukroger explores these interconnections between early modern theories of the mind and attempts to reform the philosophical agenda more generally in his discussion of the problem of how to reconcile, or choose between, the competing demands of reason and sensibility. By tracing the root of this problem back to two broader developments among the French Lockceans, on the one hand, and Hume, on the other, Gaukroger reveals how something that at first looks narrowly epistemological turns out to have a central moral and social dimension. In his analysis of Baumgarten's and Herder's strategies to counter a dilemma to which Hume's radicalization of the dichotomy between sensibility and reason had led, it furthermore becomes clear that a heightened interest in the role of sensibility called for a transformation of the style of philosophy in general.

Anik Waldow's paper ventures further into the debate about the aims and goals of eighteenth century philosophy by discussing Herder's concept of *Bildung*. Herder, she argues, effectively undermines the dichotomy between reason and sensibility by urging that individuals ought to be educated in the art of using their various affective and rational capacities in an integrated way. The use of pre-rational capacities here becomes a *sine qua non* for a successful intellectual practice. Against this background, it emerges that the push against speculative metaphysics and towards an experience-focused philosophy did not result in something that could be termed a form of reductive verificationism. Instead, it prepared the way for a diversification of those cognitive elements allowed to play a role in the formation of knowledge.

With the move away from learning, logic and metaphysics and towards a philosophical culture that recognized the value of sentiments and affective body-induced states for the cultivation of the mind, a new interest in the role of education was sparked. In Annette Pierdziwol's analysis of Rousseau's *Emile*, this education takes the form of an artful stimulation of bodily sensibilities. By contrasting this sentimental education with the educational effects of an untutored exposure to society, she demonstrates how considerations about the way the body works and affects the mind can motivate considerations about the norms of a commendable educational program. Nature here emerges as circumscribing the norms of moral education in a twofold way: on the one hand, it offers the resources with which we have to work; on the other hand, it delimits the range of skills and attributes that are to be cultivated, because only that which is deemed natural is seen as beneficial to an agent's moral development, even though, as Pierdziwol argues, what is natural for Rousseau can only be discovered in hindsight, that is, after an individual's education has been completed.

To acknowledge the body as an instrument in the cultivation of morality is of course not to deny that it is due to this body that we perceive ourselves as integral parts of the great chain of cause and effect that determines processes in nature. But although we might think of ourselves as partly determined by events in the natural world, the question emerges as to how the fact that we have understanding and are susceptible to the demands of reason and morality serves to open up a unique way of dealing with the constraints nature imposes on us.

Eric Schliesser explores this question in relation to Smith's concept of the *piacular*, that is, a form of shame felt when we become the cause of unintended harm. Schliesser argues that, according to Smith, humanity requires us to accept that we participate in the causal chain of life and therefore become liable to harms we did not intend to cause, while also acknowledging that Smith stresses the importance of highly regulated norms of atonement. Superstition is here analyzed as an element that taints the norms through which the *piacular* feeling can be discharged. Schliesser's analysis thus reiterates a point that has surfaced in many of the previous discussions, namely that for many early modern thinkers an understanding of the norms that govern our epistemic and moral practices requires that we become aware of, and acknowledge as a matter of fact, the non-rational elements that crucially determine the way in which we experience the world.

These analyses bring into focus the way in which early modern philosophers conceived of nature and norms as influencing one another in the human attempt to rationalize their epistemic and moral engagement with the world. Such a focus not only puts us in a better position to reevaluate the contemporary practice of drawing a firm line between nature and normativity, it can also help us see that today's constellation of problems might make it necessary for us to rethink the relation between the realms of nature and the human in yet another way. Michael Hampe takes up this line of thought when arguing that the ecological crisis requires us to overcome anthropocentric approaches to nature, which have their roots in the seventeenth century's humble acceptance that the human perspective is the only one available to us. By claiming that contemporary adaptations of early modern panpsychism can help us to correct this anthropocentric paradigm, he stresses that even today there is a need for us to think of philosophy as an intellectual exercise with genuinely pragmatic consequences. He thus returns us in conclusion to a thought that has prominently figured in the early modern attempt to conceive of philosophy as a form of education.

References

- Brandom, Robert. 1994. *Making it explicit*. Cambridge, MA: Harvard University Press.
- Brandom, Robert. 2002. *Tales of the mighty dead*. Cambridge, MA: Harvard University Press.
- De Caro, Mario, and David Macarthur. 2010. *Naturalism and normativity*. Chicago: Chicago University Press.
- Fodor, Jerry. 1987. *Psychosemantics. The problem of meaning in the philosophy of mind*. Cambridge, MA: MIT.
- Foot, Philippa. 2003. *Natural goodness*. Oxford: Oxford University Press.

- Haakonsen, Knud. 1998. Divine/natural law theories in ethics. In *The Cambridge history of seventeenth-century philosophy*, ed. Daniel Garber and Michael Ayers, 1317–1357. Cambridge: Cambridge University Press.
- Hatfield, Gary. 2005. The history of philosophy as philosophy. In *Analytic philosophy and history of philosophy*, ed. Tom Sorell and John Rogers, 83–128. Oxford: Oxford University Press.
- Lenz, Martin. 2010. *Lockes Sprachkonzeption*. Berlin/New York: Walter de Gruyter.
- McDowell, John. 1986. *Mind and world*. Cambridge, MA: Harvard University Press.
- Millikan, Ruth. 2004. *Varieties of meaning*. Cambridge, MA: MIT.
- Pasnau, Robert. 2011. *Metaphysical themes 1274–1671*. Oxford: Oxford University Press.
- Putnam, Hilary. 2002. *The collapse of the fact/value dichotomy and other essays*. Cambridge, MA: Harvard University Press.
- Rorty, Richard. 1979. *Philosophy and the mirror of nature*. Princeton/Oxford: Princeton University Press.
- Spinoza, Baruch de. 1994. *A Spinoza reader*, ed. Edwin Curley. Princeton: Princeton University Press.
- Stemmer, Peter. 2008. *Normativität. Eine ontologische Untersuchung*. Berlin/New York: Walter de Gruyter.
- Tartaglia, James. 2007. *Rorty and the mirror of nature*. London: Routledge.

Part I
Nature's Influence on the Mind

Chapter 2

Intentionality Bifurcated: A Lesson from Early Modern Philosophy?

Lionel Shapiro

2.1 Introduction

In the long history of philosophical reflection about intentionality—how thought is directed at the world—the seventeenth century is generally regarded as a heyday of innovation. The novelty has sometimes been exaggerated: scholars are increasingly emphasizing close connections between Early Modern theories of “ideas” and late Scholastic accounts of “concepts.”¹ In this paper, nevertheless, I will call attention to what I think is one genuinely new feature of some Early Modern theorizing about intentionality, a feature that has been overlooked. This is a pity, since it is a feature that goes missing in present-day theories, but deserves to be taken seriously.

According to the Scholastic tradition, cognition is (very roughly) a process in which the same forms that structure things in the world come to structure or “inform” our various cognitive capacities, starting with the sensory faculty and ending with the intellect. This theory is meant to accomplish two explanatory tasks at the same time. On the one hand, it is meant to address question (1).

1. What accounts for our ability to entertain a thought *that thus-and-so*, and hence to make a judgment that is true or false depending on how things are?

On the other hand, the sharing of forms between things in the world and our cognitive capacities is meant to explain our *successful use of these capacities*. If we can understand how something’s form comes to structure the intellect, this promises to let us understand how we might approach a grasp of that thing’s nature. And such a grasp would count, in the highest degree, as successful use of our cognitive capacities. So the sharing of forms is also meant to address question (2).

¹ See e.g. Perler (1996, 2010), Pasnau (1997), Hoffman (2002), Brown (2007), Clemenson (2007) and Carriero (2009).

L. Shapiro (✉)

Philosophy Department, University of Connecticut, Storrs, CT 06269, USA
e-mail: lionel.shapiro@uconn.edu

2. What accounts for the success of our proper methods of inquiry?

When Early Modern philosophers either jettison or radically reconstruct the Scholastic framework for understanding cognition, the possibility emerges that each of these explanatory tasks may require its own conceptual tools.

This paper will examine how two very different Early Modern thinkers, Descartes and Locke, find themselves needing to appeal to two different ways the same “idea” can be directed at an object. According to both philosophers, the same idea can be directed at *two different objects*. It can be the ‘idea of’ two different objects, in two different senses of the phrase.

One of these senses corresponds to the first of the above explanatory tasks: accounting for how we can make judgments that are true or false depending on how things are. Since this sense will be elucidated in terms of how ideas contribute to propositional mental acts, I will call it the *propositional* sense of ‘idea of’. In addition to the propositional sense, which is common to Descartes and Locke, I will argue that each philosopher employs his own contrasting sense. These other senses correspond to the different ways Descartes and Locke approach the second explanatory task, that of accounting for cognitive success. But since both philosophers use the word ‘represent’ in this connection, I will speak in both cases of a *representational* sense of ‘idea of’.² The structure of my interpretation is summarized in Table 2.1.

For each philosopher, the twofold intentional directedness underwrites two kinds of *normativity*. On the one hand, Descartes and Locke share the traditional view that judgments are subject to a *norm of truth*, and we will see that what it takes for a judgment to achieve truth depends on the propositional intentionality of the ideas employed in the judgment. On the other hand, both philosophers hold that there are *norms governing the improvement of ideas*, the achievement of which will be explained in terms of their representational intentionality. For both philosophers, these representational norms will be a matter of how our ideas are beholden to mind-independent *natures*: Cartesian “true and immutable natures” or Lockean “real essences.”

I have three aims in telling this story. First, I hope to show how my reading of Descartes and Locke resolves puzzles posed by their texts: in Descartes’s case the notorious puzzle of “materially false ideas,” in Locke’s case the analogous but neglected puzzle of “inadequate ideas.” (In each case, the puzzle is why what is alleged to be a deficient idea of some object doesn’t count instead as an idea of another object.) Here I will draw on interpretations I have elaborated and defended in far greater detail elsewhere.³ Second, by pointing to the very different pressures that explain why each philosopher needs to appeal to two kinds of intentionality, I hope to illuminate their respective fundamental commitments concerning cognition. In each case, the need results from a characteristic *rationalist* commitment: in

²I don’t mean to claim that these are the *only* contexts in which Descartes or Locke use what we might regard as notions of representation. For general discussion of representation in Descartes and Locke respectively, see Simmons ([forthcoming](#)) and Lennon (2007).

³Shapiro (2012, 2010).

Table 2.1 Overview

To account for	Descartes appeals to	Locke appeals to
Truth-evaluable propositional acts	Propositional intentionality	Propositional intentionality
Certain kinds of cognitive success	Representational intentionality (Cartesian version)	Representational intentionality (Lockean version)

Locke's case the rationalism concerns question (1), while in Descartes's case it concerns question (2). Third, I will suggest that understanding why Descartes and Locke are driven to bifurcate intentionality should carry a lesson even for those present-day theorists who reject both rationalist commitments.

2.2 Descartes

My first task will be to explain the propositional sense of the phrase 'idea of', by exhibiting it as a familiar sense Descartes takes for granted when elucidating his talk of ideas. After that, I will present evidence that the propositional sense doesn't suffice for his explanatory purposes, and offer an account of the contrasting representational sense.

2.2.1 *Propositional Ofness*

In a letter written to Mersenne as the *Meditations* awaited publication (AT 3:391–7, CSMK 3:184–7⁴), Descartes stresses an ordinary, theoretically noncommittal sense in which a mental entity can count as an "idea of" something. He does so in response to an anonymous objector who claims to be puzzled by his talk of "the idea of God, the idea of the soul, and (in general) the ideas of imperceptible things" from which we receive no sensory images (AT 3:375–6).⁵ Granting that by 'idea' Descartes *doesn't* mean an image retained in the imagination, the objector asks for a "clearer explanation" of what he means by these phrases. Descartes protests:

Is it possible that [the objector] could not, as he says, understand what I mean by the idea of God, the idea of the soul, and the ideas of imperceptible things? I mean only what he must necessarily have understood himself when he wrote to you that he did not understand my meaning.

⁴References in this format will be to the edition of Adam and Tannery (1964–1976, henceforth AT), and to the translations in Cottingham et al. (1984–1991, henceforth CSM and CSMK). Where I have significantly modified a translation, the original is provided in brackets.

⁵Mori (2012) offers biographical, textual and philosophical evidence that the objector, whose letter Descartes received from Mersenne, was Hobbes.

He goes on to explain:

For he does not say that he conceived nothing by [*n'ait rien conceu par*] the expressions 'God', 'soul', 'imperceptible things'; he just says that he did not know what was to be understood by the *idea of these things*. But if he conceived anything by these expressions, as he doubtless did, he knew at the same time what was to be understood by the ideas—namely nothing other than the very thing [*cela mesme*] he conceived. (AT 3:392, CSMK 3:184, emphasis added)

The point here is a simple one. My idea of God, Descartes is saying, is whatever I conceive when I use the word 'God', and my idea of the soul is whatever I conceive when I use the word 'soul'. He then generalizes over all such examples:

[W]e cannot express anything by our words, when we understand what we are saying, without its being certain thereby that we have in us the *idea of the thing* which is signified by our words. (AT 3:393, CSMK 3:185, emphasis added; see also the Second Replies at AT 7:160, CSM 2:113)

If I conceive anything at all in using the word 'God', Descartes is saying, my mind has *the idea of God*. If I conceive anything at all in using the word 'soul', my mind has *the idea of a soul*. And so on.⁶

When explaining what he means by an "idea of something," why does Descartes point to a relation between ideas and *words*? The answer is that his explanation in terms of words follows from a more fundamental principle that doesn't involve words. According to him, ideas serve to determine the overall "form" of "thoughts" (AT 7:160, CSM 2:113; AT 7:37, CSM 2:25–6). For example, it is in virtue of the ideas it involves that one thought is a thought *that God is a deceiver* and another thought is a thought *that souls are extended*. I can now state the fundamental principle Descartes is relying on:

Proposition principle An *idea of x* is an idea one can employ (as the relevant component of one's thought) in entertaining the thought *that ... x...*

For example, a person's idea of God is the one she employs when she entertains the thoughts *that God is a deceiver*, *that she is not God*, etc. This explains Descartes's next move in response to the objector: "how could he have said that God is infinite and incomprehensible ... unless he had an idea of him?" (see also AT 7:188, CSM 2:132; AT 9A:210, CSM 2:273).⁷

We can now understand why Descartes answers his critic by pointing to a connection between ideas and words. Suppose that I understand what I'm saying in using the words 'God is a deceiver'. Then I am entertaining the thought *that*

⁶Cf. Perler (1996, 13–14).

⁷It may be objected that in the same letter Descartes rejects a distinction between two kinds of ideas, those "expressed by terms" and those expressed "by propositions" (AT 3:396, CSMK 3:186). For my purposes, however, it won't matter whether we talk about a thought *that God is infinite* or a thought *of God as infinite*, as long as this specification immediately determines a truth condition for a possible judgment. Cf. Perler (1996, 257–61) and Nuchelmans (1983, 42–4). For the dialectical context of Descartes's rejection of the above distinction, see Mori (2012).

God is a deceiver. According to the proposition principle, the idea that is the element of this thought corresponding to my word ‘God’ must be the idea *of God*. The principle also explains a connection between what an idea is “of” and the truth conditions for judgments that involve the idea. If one of your ideas is *of God* and a second idea is *of infinity*, you can employ them in an affirmation that is true just in case God is infinite.

Whenever an idea is “of” an object in the sense Descartes has explained, and which I have explicated using the proposition principle, let us say that the idea is *of that object in the propositional sense*.⁸ Martin Lenz calls attention to the importance of this sense for Early Modern theorists of ideas:

For many Early Modern authors, ideas are the components of beliefs that can be expressed using sentences. For example, if I have the belief that milk is white, I therefore also have at my disposal the ideas of milk and of white.⁹

But the propositional sense of ‘idea of’ should also be familiar from present-day discussions of concepts. Consider, for example, Jerry Fodor’s influential capital-letter notation. Where Early Modern theorists would speak of the ‘idea of an animal’, Fodor speaks of ‘the concept ANIMAL’. He explains: “[C]oncepts are constituents of mental states. Thus, for example, believing that cats are animals is a paradigmatic mental state, and the concept ANIMAL is a constituent of the belief that *cats are animals* (and of the belief that *animals sometimes bite*, etc....).”¹⁰ Accordingly, the concept ANIMAL counts as being the concept of an animal in the propositional sense. Notice that neither Descartes nor Fodor needs to appeal to any *relation of representation* to elucidate what they mean by the idea of God or the concept ANIMAL.

2.2.2 Why Propositional Ofness Is Not Enough

I will now argue that Descartes insists that what an idea is of in the propositional sense can *diverge* from what the same idea is of in some second sense, for which he uses the language of “representation.” My evidence will come from his much-discussed reply to an objection by Arnauld concerning “material falsity.”¹¹ Unlike

⁸ Elsewhere, I have argued that this sense is essential to understanding one strand of Descartes’s epistemology: his view that we can delineate the natures of things by contemplation of what is contained in our clear and distinct *ideas of them* (Shapiro 2012, 396–8).

⁹ Lenz (2010a, 254, my translation).

¹⁰ Fodor (1998, 6).

¹¹ The extensive literature on Descartes and material falsity is examined in Shapiro (2012); see also De Rosa (2010) and Naaman-Zauderer (2010). My reading of Descartes’s reply to Arnauld is most influenced by Margaret Wilson, who proposes that Descartes is distinguishing between two “senses of ... ‘idea of’,” namely a “presentational” and a “referential” sense (Wilson 1990, 69–70, 73–4). But neither of Wilson’s two senses lines up with *either* of the two senses I attribute to Descartes (see Shapiro 2012, 401).

most readers,¹² I will take Descartes to be making a fundamental point about intentionality, rather than one specifically about *sensory* representation.

In the Third Meditation, Descartes's meditator makes a remark that puzzles Arnauld. He says that while there is a sense in which only judgments can be false, "there is a certain other falsity, a material kind, which occurs in ideas [*quaedam alia falsitas materialis in ideis*], when they represent non-things as things" (AT 7:43, CSM 2:30). For example, the meditator explains, "if it is true that cold is nothing but the privation [*privatio*] of heat, the idea which represents it to me as something real and positive deserves to be called false...." Arnauld interprets the meditator's scenario as follows.

Third Meditation scenario My *idea of cold* might turn out to represent something other than cold to me. In that case it would misrepresent its object, and therefore count as a *false idea of cold*.

Arnauld then objects that this scenario is incoherent. According to him, an idea that represents something other than cold "is not the idea of cold" (AT 7:207, CSM 2:146). That is why, "if cold is merely a privation, there cannot be an idea of cold which represents it to me as a positive thing..." (AT 7: 206, CSM 2:145).

In his reply, Descartes insists that the Third Meditation scenario is coherent, as long as its words are understood as he had intended.¹³ Arnauld's points, he writes,

can be readily accepted if they are taken as he himself intends them [*facile possunt admitti, prout ipse illa intelligit*]. But I meant what I wrote in another sense [*quae scripseram, alio sensu intellexi*], which seems to me to be equally correct.

He will thus "dodge [Arnauld's] blows" by distinguishing senses of words. Let us see if we can identify the senses he has in mind. Recall the key premise of Arnauld's objection: an *idea of cold* is an idea that represents cold to the mind. This is how Descartes evaluates Arnauld's premise:

When [Arnauld] says 'if cold is merely a privation, there cannot be an *idea of cold* which represents it as a positive thing', it is clear that he is dealing solely with an idea taken in the *formal* sense... [W]hen we think of [ideas] as representing something we are taking them not *materially* but *formally*. (AT 7:232, CSM 2:162–3)

Here he is invoking common late Scholastic terminology, found in Suárez as well as the Coimbra commentators he studied at school.¹⁴ To take a concept or idea *formally* is to regard it insofar as it *represents* some object to the mind. Expressed using vocabulary Descartes uses interchangeably, it is to consider that *object* insofar

¹² An exception is Wee (2006, 55–8).

¹³ For criticism of common readings according to which Descartes *agrees* with Arnauld that the scenario is incoherent, and denies that it is what he had in mind in the Third Meditation, see Shapiro (2012, 391–3).

¹⁴ Suárez (1597/1960–1966, 8.3.16/2:106) is cited by Wells (1984, 32n36). For similar passages from Pedro da Fonseca and Antonio Rubio, see Clemenson (2007, 43–5).

as it “is objectively in the intellect.”¹⁵ Descartes concedes that Arnauld’s premise is true if the term ‘idea of cold’ is taken in the “formal sense” Arnauld intends. By contrast, he says, Arnauld’s premise is false if ‘idea of cold’ is taken in Descartes’s own intended sense, which is apparently the above-mentioned “material” sense:

[I]f cold is simply a privation, the idea of cold [taken materially] is not cold itself insofar as it is objectively in the intellect [that is to say, it does not represent cold], but something else, ... namely a sensation which has no existence outside the intellect. (AT 7:233, CSM 2:163)

Taken in the material sense, the idea is still *of cold* (here supposed to be a privation), even though what the idea *represents* to the mind is a mere sensation.¹⁶ Hence the Third Meditation scenario, interpreted as Descartes had intended, is perfectly coherent.

But what is Descartes’s intended “material” sense, such that whether an idea is “of cold” in this sense doesn’t depend on its representing cold to the mind? I will now argue that *the material sense is simply the propositional sense*. That is because this reading makes best sense of how he explains the consequences of taking the idea of cold materially:

[I]f we were taking them materially, [ideas] would have no regard to [*nullo modo ... respicerent*] the truth or falsity of their objects ... Thus, whether cold is a positive thing or a privation, I do not on that account have a different idea of it [taken materially, that is], but there remains in me the same one I have always had [*non aliam idcirco de ipso habeo ideam, sed manet in me eadem illa quam semper habui*] (AT 7:232, CSM 2:163).

His point seems to be this. Which of my ideas counts as an *idea of cold in the propositional sense* doesn’t depend on whether cold turns out to be something metaphysically “true” or “false”—a real quality or an unreal privation. By contrast, which of my ideas counts as *representing cold to my mind* does depend on what cold turns out to be. If cold is a positive being, a “different idea” of mine will count as being “of cold” in the *formal sense* than if cold is a privation. So, depending on what cold turns out to be, I have a “different idea of it” in the formal sense. But whether cold is a positive being or a privation, I certainly have “the same idea of it” in the propositional sense. This is the idea I express by ‘cold’, the idea I use when I wonder, e.g., whether cold might be a privation. Taken in the propositional sense, it can’t turn out that *this idea* isn’t really “of cold”!

We have now found Descartes’s explanation of why the Third Meditation scenario is coherent. The idea of mine that is guaranteed to be *of cold* in the “material” or propositional sense might turn out to be an *idea of something else* in the “formal sense.”

¹⁵ More precisely, I claim that Descartes uses the two locutions interchangeably when discussing the intentional directedness of ideas. I take no position on the contested *ontology* of objective being in Descartes: i.e. whether or not the act of the intellect that represents some object is really distinct from the object insofar as it is objectively in the intellect, what Descartes elsewhere call an idea “taken objectively” (AT 7:8, CSM 2:7). See e.g. Ayers (1998, 1067–8).

¹⁶ I defend this identification of what the idea in Descartes’s example represents in Shapiro (2012, 389–90). For a contrary position, see De Rosa (2010, 55, 69–79).

2.2.3 *Representational Ofness*

Where Descartes speaks of the idea of an object in the “formal sense,” let me speak of an *idea of that object in the representational sense*. So far, I have said nothing to explain this sense, but have merely argued that Descartes distinguishes representational from propositional ofness. To understand the role that representational ofness plays in Descartes’s thought, it will be useful to have a few examples.

The Third Meditation scenario under discussion between Descartes and Arnauld is purely hypothetical: there is reason to think Descartes doesn’t regard cold as a privation.¹⁷ But he says that the same predicament as in his hypothetical case “often happens in the case of obscure and confused ideas...” (AT 7:233, CSM 2:163). What actual examples, then, might he have in mind of ideas that are *propositionally* of something they fail to be *representationally* of? Consider first the confused idea a non-philosopher expresses by the word ‘cold’.¹⁸ Perhaps this idea is *propositionally* of a quality in bodies, a quality the vulgar call ‘cold’, one that from the standpoint of Descartes’s science has turned out to be a certain state of motion. Yet, as I have interpreted the Fourth Replies, the same idea is *representationally* of another object, namely a sensation which exists only in the mind. The non-philosopher may confuse this sensation with the object the idea is *propositionally* of, namely (we are supposing) the state of motion she calls ‘cold’. This would explain why she is liable to form a false judgment as to the latter quality’s nature.¹⁹

Second, there is the Third Meditation’s example of the ordinary “idea of the sun” which seems to be “acquired from the senses” (AT 7:39, CSM 2:27). This is an example used by the Coimbra commentators, who argue that the perception represents a small disk, not the sun.²⁰ I believe there is reason to hold that Descartes is *agreeing* with them that this idea *represents* a small disk, rather than the sun.²¹ Still, on his view, it remains an *idea of the sun in the propositional sense*. It is the idea a peasant uses in forming the false judgment that the sun is smaller than the earth. Indeed, what accounts for this false judgment is that the peasant confuses the two objects the idea is “of” in the two senses.

¹⁷ Hoffman (1996) and Nelson (1996).

¹⁸ For a recent discussion that includes a survey of the literature on Cartesian sensory ideas, see De Rosa (2010).

¹⁹ Compare Descartes’s description of the hypothetical materially false idea: “the obscurity of the idea gives me occasion to judge that this idea of the sensation of cold represents some object called ‘cold’ which is located outside me” (AT 7:234–5, CSM 2:164). Though the idea is *in fact* representationally of a sensation, I may take it to be representationally of its *propositional* object, namely the quality I called ‘cold’. Even if this mistake doesn’t amount to an explicit judgment (one only a philosopher could make!), it can account for my false judgment about that quality.

²⁰ In Doyle (2001, 162–3 and 172–3); see also Wells (2003, 27–30).

²¹ Here I agree with Clemenson (2007, 71). See also Grüne (2010, 34–5n38), who sees this consequence of her reading as a difficulty.

Table 2.2 Cases where propositional and representational ofness may diverge

	Is propositionally of	Is representationally of
Idea the vulgar express by ‘cold’	A quality in bodies (which turns out to be a state of motion)	A sensation in the mind
Ordinary sensory idea expressed by ‘the sun’	The sun	A small disk
Idea a pagan priest expressed by ‘ <i>Iuppiter</i> ’	Jupiter	The true God

A third likely example comes from a letter written in 1645:

When the ancients named [*nommoient*] many gods they did not mean many all-powerful gods, but only many very powerful gods, above whom they imagined a single Jupiter as sovereign; and consequently, to this Jupiter alone they applied the idea of the true God, this idea being presented to them in a confused manner. (AT 4:188, CSMK 3:248)

Here Descartes appears to be saying that the idea a pagan priest expressed by ‘*Iuppiter*’ is in one sense an idea of the true God.²² Clearly, however, the pagan’s idea is propositionally of Jupiter. When the pagan asserts the Latin equivalent of ‘Jupiter is subject to the fates’ (cf. AT 1:145, CSMK 3:23), he is asserting that Jupiter is subject to the fates, not that the true God is subject to the fates. The pagan confuses the true God his idea represents with the mythological object the idea is propositionally of.

These potential examples, summarized in Table 2.2, may help us answer the most pressing question: Why does Descartes appeal to representational ofness in addition to propositional ofness? My proposal is that the role of representational ofness needs to be understood in terms of Descartes’s theory of clear and distinct ideas. On one attractive reading, a clear and distinct idea is an idea that perspicuously *reveals the nature* of the object it is an “idea of” in the propositional sense. For example, a clear and distinct idea of body is an idea that enables one to judge with certainty that body is extended substance.²³ According to Descartes, rational reflection of the sort exemplified by the *Meditations* allows us to progress from obscure or confused ideas to clear and distinct ideas. What object is the resulting idea *propositionally of*, what object’s nature does it perspicuously reveal? I propose that this is the very object the original idea already *represented to the mind*.

Reflective improvement of ideas If *IDEA1* is *representationally of* *x*, then there is a process of rational reflection that takes the mind from *IDEA1* to some *IDEA2* that (i) is *propositionally of* *x*, and (ii) counts as *clear and distinct* in the sense that it perspicuously reveals the nature of *x*.

²² See Wee (2006, 56–7).

²³ For similar readings of clarity and distinctness see Alanen (2003, 159) and Clemenson (2007, 72–5).

Table 2.3 A case of Cartesian idea-improvement

	Is propositionally of	Is representationally of
<i>IDEA1</i> (confused idea a pagan priest expresses by ‘ <i>Iuppiter</i> ’)	Jupiter	The true God
<i>IDEA2</i> (clear and distinct idea resulting from reflection)	The true God	The true God

Take, for example, the confused idea our pagan priest expresses by ‘*Iuppiter*’ (see Table 2.3). While this idea is *propositionally of* Jupiter, it is *representationally of* the true God. This status explains how the pagan is capable of reflectively refining the idea to yield a clear and distinct idea that is *propositionally of* the true God, and perspicuously reveals his nature. And this is the sense in which “the idea of God is imprinted on the human mind in such a way that everyone has within himself the power to know him” (AT 4:187, CSMK 3:248).

It is this kind of presence of an object to the mind that representational ofness consists in: a kind of presence for which Descartes is happy to borrow the Scholastic terminology of “objective being in the intellect.” My thesis so far has been that this kind of Cartesian intentionality, a kind responsible for our ability to achieve clear and distinct understanding of “true and immutable natures,” needs to be distinguished from the kind I have called propositional, the kind responsible for a judgment’s truth conditions.

2.3 Locke

I will now argue that Locke too finds himself needing to invoke a kind of ofness that contrasts with propositional ofness.²⁴ One place in the *Essay* where he does so is when discussing our complex ideas of the species of substances. These are the ideas expressed by words like ‘water’, ‘gold’, ‘dog’, ‘man’, etc. My proposal is that what Locke explains in terms of the doctrine of the “nominal” and “real essences” of *substance species* can be illuminated in terms, respectively, of the propositional and representational intentionality of *substance ideas*.²⁵

²⁴ See Lenz (2010b, 345ff; 2010a, 271–2) for a related attribution to Locke of two dimensions of content, which likewise stresses the role of Locke’s ideas in mental propositions.

²⁵ In Shapiro (1999, 587–90), I proposed that Locke’s distinction between the nominal and real *essences* of substance kinds coincides with a distinction between kinds of *content* (cf. Prinz 2000). Here, I follow Shapiro (2010) in explaining that distinction in terms of propositional and representational intentionality.

2.3.1 *Propositional Ofness*

Let me first use an example to illustrate Locke's appeal to *propositional* ofness when discussing substance-ideas. Locke quotes from a recently published collection of anecdotes:

When the Abbot of St. Martin was born, ... he had so little the Figure of a Man, that it bespake him rather a Monster. 'Twas for some time under Deliberation, whether he should be baptized or no. However, he was baptized and declared a Man provisionally.... (3.6.26:454²⁶)

Suppose we wish to determine whether this oddly-shaped fellow is a man, i.e. whether he deserves to be “ranked” into the human species—or, in Locke's preferred linguistic formulation, whether he has a “right to the name *Man*” and thus has the “nominal essence” of this sort or species (3.3.12:414–5, 3.3.18:419).²⁷ What we're really interested in when we ask this question is the truth of a “mental proposition,” namely the proposition *that an animal with such-and-such features and capabilities is a man*. According to Locke's theory of “mental propositions,” this proposition is formed by “joining” two ideas (4.5.2:574, 4.5.5:575–6).²⁸ What are these two ideas “of”? As long as ‘idea of’ is used in the propositional sense, there is an immediate answer: the first is the idea *of such-and-such an animal*, and the second is the idea *of a man*. Finally, the proposition will be true provided these two ideas “agree” with each other with respect to “necessary connexion” in the same subject (4.1.3:525, 4.7.5:594). And that requires that whatever “conforms to” the idea *of such-and-such an animal* necessarily also “conforms to” the idea *of a man*.²⁹

Summarizing, we can say that what it takes for something to *belong to the species man* is for it to conform to an idea that is “of a man” in the *propositional sense*, such as the idea I express by ‘man’ (3.3.12:414–5, 3.6.7:443). Suppose now that my idea is the one Locke says people “commonly” express by that word, namely the idea of “a Body of the ordinary shape, with Sense, voluntary Motion, and Reason

²⁶ References in this format will be to Locke's *Essay* (by book, chapter, section, and page number in [Locke 1975](#)).

²⁷ Though Locke insists that his usage of ‘species’ and ‘sort’ is a mere matter of Latin versus English (3.3.12:414, 3.5.9:434), ‘species’ typically expresses a contrast with ‘genus’ less often expressed by ‘sort’ (e.g. not at 3.3.15:417, 3.6.1:439). Locke's English for ‘genus’ is ‘kind’ (2.32.6:386, 3.1.6:404), though this term too is often used without any intended contrast (2.32.24:393, 4.6.4:580). Here I will use ‘species’ as the Lockean term.

²⁸ Perler (1996, 261n) notes that Locke's theory of mental proposition has no counterpart in Descartes. This difference needn't prevent us from isolating a common notion of propositional ofness in both philosophers: see note 7 above.

²⁹ Ott (2012, 1087), stresses the connection between the truth of a Lockean mental proposition (the agreement of its component ideas) and what the ideas in the proposition are of: “whether ideas agree or disagree is a function of ... what they are ideas *of*.” Though Ott's discussion concerns simple ideas, his “*of*-ness” thus plays the role of what I am calling propositional ofness.

join'd to it" (4.6.15:589–90). Then I can say that being a man is a matter of being a body of the ordinary shape with sense, voluntary motion and reason joined to it. In short, by contemplating the composition of a complex idea of mine that counts—in the *propositional* sense—as an idea of man, I can come to know necessary and sufficient conditions for *being a man*. This will be important in Sect. 2.4.

2.3.2 Why Propositional Ofness Is Not Enough

I am now ready to argue that Locke also appeals to a second, *non-propositional* sense in which an idea can be “of” a species of substance. Switching from *man* to Locke’s other favorite species, suppose that the complex idea we express by the word ‘gold’ is composed out of ideas of *yellowness*, *a certain density*, *malleability*, *fusibility*, *solidity*, and *extension*, together with the idea of their union “in an unknown *substratum*” (cf. 2.24.37:317, 2.32.23:392 and 3.6.2:439). Call this complex idea *IDEA1*. Locke stresses that *IDEA1* fails to be “*a perfect complete Idea, of a sort of things*” (2.32.23:392). For it omits, among other features, *fixedness* (non-volatility), which has been observed to coexist with the features already included. And he urges natural historians to “rectify” this imperfection:

[W]e ... must enquire into the Nature and Properties of the Things themselves, and thereby perfect, as much as we can, our *Ideas* of their distinct Species ... [W]e must, by acquainting ourselves with the History of that sort of Things, rectify and settle our complex *Idea*, belonging to each specifick name. (3.11.24:520–1; see also 3.11.25:521–2)

Suppose, in what follows, that we have followed Locke’s advice and added the idea of fixedness to *IDEA1*, and that we now use our word ‘gold’ to express the resulting “perfected” idea *IDEA2*.

We can now formulate a puzzle modeled on Arnauld’s puzzle for Descartes. As we will soon see, Locke repeatedly makes two claims that appear to be inconsistent:

- (a) *IDEA1* and *IDEA2* are ideas of *distinct species*.
- (b) *IDEA2* is a more perfect idea of *the same species* of which *IDEA1* was an idea.

Notice the parallel with Arnauld’s charge that Descartes is committed to the claim that a given idea might simultaneously be a “materially false idea of cold” and an idea of *something other than cold*. Unlike Arnauld’s puzzle, this puzzle hasn’t received attention; indeed, Locke’s claim (b) has been overlooked. Several readers have tried to explain what he means when he speaks of *more or less perfect* substance ideas.³⁰ What hasn’t been appreciated is that he speaks of more or less perfect ideas of *the same species of substance*.

Once again, the contradiction would disappear if we concluded that ‘of’ is used in different senses in (a) and (b). This already gives us reason to suspect that Locke

³⁰ See esp. Ayers (1991, vol. 2: 75–7), Mattern (1986) and Atherton (2007).

Table 2.4 A case of Lockean idea-improvement

	Is propositionally of	Is representationally of
<i>IDEA1</i> (idea I formerly expressed by ‘gold’)	A species distinct from <i>gold</i> (doesn’t require fixedness)	[Same as below]
<i>IDEA2</i> (perfected idea I now express by ‘gold’)	The species <i>gold</i>	Call this species <i>G</i> (it is distinct from <i>gold</i>)

may be working with two kinds of intentionality. The key to distinguishing and identifying them is to notice that claims (a) and (b) occur in entirely different contexts in the *Essay*. The places where Locke endorses claim (a) are all contexts where he is concerned with explaining what counts as gold, silver, a man, etc. Since “*each abstract Idea, with a name to it, makes a distinct Species,*” he argues, the “*sorting of Things is the Workmanship of Men*” (3.6.37–8:462–3; also 3.5.7:432, 3.3.14:416–17). As explained above, whether some object *counts as gold* depends for Locke on whether the object conforms to the *idea of gold* in the *propositional* sense. So ‘idea of’ in claim (a) expresses propositional-ness. This lets us fill in our first column in Table 2.4. After all, the idea I now express by ‘gold’ is, in the propositional sense, an *idea of gold*. According to Locke’s claim (a), then, the idea I *formerly* expressed by the word ‘gold’ must have been an idea of a distinct species, one that doesn’t require fixedness.

By contrast, Locke endorses claim (b) in contexts where he isn’t focusing on the role of our ideas in determining what counts as gold, silver, a man, etc. Instead, when he stresses that we “*have different Ideas of the same Substance*” (3.9.13:482–3), he is explaining that our complex ideas are “*imperfect and inadequate*” as “*representations*” of the species of substances found in nature (2.31.6:378, 2.31.11:382, 2.31.13:383).³¹ As we have seen, he concludes from this that “*we must ... perfect, as much as we can, our Ideas of their distinct Species.*” Yet he concedes that our ideas will never be “*exact Representations*” of these species (3.9.20:488; also 3.9.17:486, 2.31.8–10:380–2, 2.32.24:392–3, 4.3.26:556–7). In other words, *IDEA1* and *IDEA2* are more or less imperfect representations of the same species of substance.³² Let us refer to the species in question, the one mentioned in claim (b), as *G*. Locke himself, well aware that he can’t identify this species as gold, picks out *G* using descriptions such as “*that sort of Substance, we denote*

³¹ There is an explicit cross-reference from the opening sentence of 3.9.12:482 to the opening sentence of 2.31.6:378. In both places, Locke contrasts the inadequacy he has in mind with a second inadequacy: substance ideas also count as inadequate when (inappropriately) regarded as revealing the “*real essences*” of species.

³² In several of these passages, Locke doesn’t expressly note that he is talking about *species* rather than individuals. Yet taken together, and in their contexts, they demand the former construal. In one instance, Locke clarifies his intention in the *Essay*’s fourth edition. Rather than speak of that person who “*has the perfectest Idea of any particular Substance*” he now speaks of that person who “*has the perfectest Idea of any of the particular sorts of Substance*” (2.23.7:299).

by the word *Gold*” (2.31.9–10:381–2) or “that sort of Body the Ring on his Finger is made of” (3.9.17:486).³³

In short, I have argued that according to Locke, an idea that is “of gold” in the propositional sense will also count as an idea “of” some other species in a sense relevant to *representation*.

2.3.3 *Representational Ofness*

To understand how *IDEA1* and *IDEA2* count as representing the same species *G*, we must turn to the three chapters of the *Essay* (2.30–32) devoted to ideas taken “in reference to things ... which they may be supposed to represent.” Here Locke’s central question is this: To what extent do our various ideas “have a Conformity ... with their Archetypes”? By ‘archetype’ he means a model to which something else is *intended to conform*. And he argues that two kinds of ideas, simple ideas and complex ideas of substances, are appropriately intended to conform to archetypes outside the mind. He stresses that the relevant “conformity” between idea and object isn’t the literal “sharing of form” appealed to by the Scholastic tradition. Rather, it is merely a “steady correspondence” (2.30.2:373): what we would today call a *correlation*.

Locke’s account of *representation as conformity to an appropriate archetype* is spelled out most explicitly for simple ideas.³⁴ But there is every reason to think the general account is supposed to apply to complex ideas of substances as well. Here conformity by correlation needs to be understood as follows.

Conformity by correlation Suppose I have a complex idea (propositionally) of a substance with features F_1, F_2, \dots, F_n . My idea conforms by correlation to a species *S* to the extent that having features F_1, F_2, \dots, F_n correlates with belonging to species *S*, and can thus serve to distinguish things of that species. (For sets of qualities serving as “marks” used to “distinguish” substance species, see 2.23.8:300 and 4.6.11:585.)

³³ In two striking passages (analyzed in Shapiro 1999, 561–4), Locke complains of how “hard it is” to specify the species of which *IDEA1* and *IDEA2* are imperfect ideas without “cross[ing his] purpose” by using the word ‘gold’ rather than just mentioning it (3.6.19:449, 3.6.43:465–6).

³⁴ In Shapiro (2010), I argue that the representational ofness of *simple ideas* diverges from their propositional ofness. For a similar view, see Ott (2012, 1089–93). Ott reserves the label “representation in the strict sense—the *of*-ness of an idea” for a notion that plays the *propositional* role (see note 29 above). He distinguishes this “*of*-ness” of ideas from their “role as marks or signs” that corresponds roughly to what I call Locke’s *representational* ofness. Since Ott and I recognize that Locke uses ‘represent’ in both ways (Shapiro 2010, 579n23), this difference is chiefly terminological. There remain two substantive differences. First, Ott’s Locke regards color-ideas, *unlike shape-ideas*, as “blank effects” lacking the “*of*-ness” that corresponds to my propositional ofness. On my reading, by contrast, color- and shape-ideas each possess *both* kinds of intentionality. Second, unlike Ott, I argue that Locke takes shape-ideas to be perfect “marks or signs” of *powers to produce ideas* (Shapiro 2010, 561). Locke doesn’t, implausibly, take these ideas to be perfect marks or signs of the very shapes they are propositionally of.

I can now say what I think it means for an idea to be “of” a species of substance in Locke’s representational sense.

Representation ofness and adequacy My complex idea is *representationally* of a substance species *S* provided I appropriately intend that the idea conform by correlation to *S*. And such an idea will *adequately represent* the species *S* to the extent that it *does* conform by correlation to *S*.

This account of representation explains what Locke means when he approves of those natural historians who intend their substance ideas to be “Representations in the Mind, of Things that do exist, by *Ideas* of those qualities that are discoverable in them” (2.31.6:378, 2.31.8:380–1, 3.9.13:482). And it explains why he says, concerning our “*specific Ideas of Substances,*” that we should “endeavour ... to make them as complete as we can, whereby I mean, that we should put together as many simple Ideas, as being constantly observed to co-exist, may perfectly determine the *Species*” (4.12.14:648).

But what is the species *G* such that I appropriately intend that the combination of features singled out by the idea I express by ‘gold’ should correlate with membership in *G*? To answer the question, consider our *purpose* in employing substance ideas. Here is Locke’s explanation of our purposes in forming all “abstract complex *Ideas*”:

[T]he first Thing [the mind] does, as the Foundation of the easier enlarging its Knowledge, ... of the things themselves, that it would know ..., is to bind them into Bundles, and *rank them so into sorts, that what Knowledge it gets of any of them, it may thereby with assurance extend to all of that sort....* (2.32.6:386, emphasis added)

In the case of substance ideas, this goal will amount to *inductive projectibility*. When I discover that a particular piece of gold was dissolved by *aqua regia*, Locke holds, I can with fair “assurance” predict that other pieces will likewise be dissolved.³⁵ (For his interest in induction, see 4.3.29:560, 3.6.24:452, 4.6.13:588, and 4.12.10:645.) It is our concern for inductive projectibility, then, that underwrites Locke’s injunction that we should perfect our ideas of substance species so as to make them better represent their archetypes. Calling attention to two of the claims cited above, Richard Boyd remarks: “Officially in these passages Locke is addressing the problem of the ‘inadequacy’ of kinds [that is: ideas of kinds] of substances rather than the problem of inductive categories, but ... I am inclined to think that he means the solution to apply to the latter problem as well.”³⁶

Boyd’s suggestion, which is bolstered by the displayed passage, lets us explain how a species *G* is picked out as the one represented by the idea I express by ‘gold’.

³⁵This claim would require serious qualification to be defensible. However, here is all Locke need be committed to: the more we “perfect” a substance idea, the more we may expect that *there will be reliable generalizations* concerning “coexistence” of the features collected in the idea with additional ones, and this expectation is relevant to induction.

³⁶Boyd (1991, 131–2). If this is right, Locke’s position resembles Boyd’s own conception of “natural kinds” (1999, 2010), as well as that of Millikan (2000).

The key is that there is reason to attribute to Locke the following claim about what grounds inductive projectibility.

Projectibility and explanatory constitutions A substance idea is inductively projectible to the extent that it conforms by correlation to a species whose members share a common “internal constitution,” one that explains their possession of the features already included in the idea, *as well as those additional features one will come to include if one continues to perfect the idea (by including features that have hitherto been “constantly observed to co-exist” with the set already included).*

As Locke says, such a shared “real constitution ... is the foundation of all those Properties, that are combined in, and are constantly found to co-exist with the *nominal Essence*” (3.6.6:442). According to the Scholastic metaphysical framework he adapts to his purposes, this internal constitution serves as the “real essence” of the species in question, the source from which “flow” the “endless” additional essential “properties” or *propria* of the species (2.32.24:392–3, 2.31.10:382, 3.11.22:520, 4.6.11:585).³⁷

In short, Locke is saying that I appropriately intend that the complex idea I express by ‘gold’ should conform by correlation to the species *G* that is picked out by possession of this underlying constitution. This means that my idea is *representationally* of the species with the explanatory internal constitution. Yet my current idea is highly unlikely to *perfectly represent* that species. There will be things that agree to my idea, and hence belong to the species my idea is propositionally of (namely the species *gold*), but don’t belong to the species my idea is representationally of.

This imperfect representation accounts for failures of induction. Locke recounts the “sad Experience” made by chymists “when they, sometimes in vain, seek for the same Qualities in one parcel of Sulphur, Antimony, or Vitriol, which they have found in others” (3.6.8:443). Though the chymists intend their substance ideas to conform to sorts with explanatory internal constitutions, their ideas end up doing so only *imperfectly*. Their ideas fail to be “exact representations” of the archetypes they are intended to conform to, and thus stand in need of perfecting. As Peter Anstey has recently observed, in the course of a reading that in many respects matches the one I have summarized here and elaborated elsewhere, Locke is a “convergent conventionalist” about our ideas of substance species.³⁸

³⁷ On Locke’s use of ‘property’ to mean *proprium* in the sense derived from Porphyry, see e.g. Ayers (1991, vol. 2, 21, 67–74) and Pasnau (2011, 658–60). Contrary to Pasnau, I am arguing that Locke’s embrace of natural kinds in the *Essay* does “depend on the notion of an explanatory essence” from which a species’ observable properties flow.

³⁸ Anstey (2011, ch. 11); cf. the section “Converging on a real essence” in Shapiro (1999, 576–82). Kornblith (1993, ch. 2) sketches a similar account of Locke’s implicit view of “chemical method,” but claims it contradicts Locke’s “official position.” Several interpreters besides Anstey and myself have argued that Locke officially embraces natural kinds. However, their conceptions differ from the one I intend. According to Conn (2002), what is required for there to be natural kinds is that (a) *each set of perceptible features corresponds to some shared internal constitution.*

2.4 Cartesian and Lockean Rationalism

We have now seen both Descartes and Locke make use of a distinction between two kinds of intentionality, each of which is appealed to in explaining or describing a different aspect of our mental lives. One kind, which both philosophers address using Scholastic metaphysical frameworks, is appealed to when answering question (2) of my Introduction: What accounts for the success of our proper methods of inquiry? In Descartes's case, the method in question is that of attaining, through rational reflection, clear and distinct ideas that perspicuously reveal fundamental natures. In Locke's case, the method in question is inductive inference, success at which requires that our substance ideas conform by correlation to sorts picked out by explanatory real essences. Each philosopher finds that the kind of intentionality he invokes in answer to question (2) isn't straightforwardly relevant to question (1): What accounts for our ability to think *that such-and-so*, and hence to think *of something* in the propositional sense? Let us now compare the respective reasons why Descartes and Locke require different kinds of intentionality to answer the two questions.

The driving force behind Locke's divorce between propositional and representational ofness is his approach to question (1). Here he advocates a rationalist thesis.

Lockean rationalism Reflection on what is required for thinking *of gold* or *of a man* (in the propositional sense) reveals informative necessary and sufficient conditions for something to count as gold or count as a man.³⁹

This rationalism is tied to Locke's preoccupation with the possibility of clarifying and resolving apparent disagreements, such as the one concerning the abbot of St. Martin. Once you convey to me the simple components of the idea you express by 'man', we can agree on informative truth conditions for the propositions you express using that word (cf. 3.9.16:484–5, 3.11.7:511–12). Yet, quite reasonably, Locke concedes that the results of such reflection have no obvious bearing on the question of what structures in the world our successful use of such thoughts depends on.

But Descartes's example shows that a divorce between propositional and representational ofness needn't have this motivation. Unlike Locke, Descartes has remarkably little to say about question (1). There is no Cartesian theory of propositional

As I read Locke, he denies (a). Instead, he holds there to be natural kinds in the sense that (b) *some ways of sorting objects by their perceptible features count as more natural than others, by coming closer to corresponding to a shared internal constitution*. Though Stuart (1999, 285–91) discusses Boyd and Kornblith, I read him as attributing to Locke a qualified version of (a), not (b). On the other hand, Pasnau (2011, 642–7) attributes an embrace of natural kinds in a very strong sense, for which I find no evidence in the *Essay*: “there is a unique system of species (and higher genera) that best captures the similarities and differences among individuals.” For an analysis of Locke's carefully reasoned ambivalence about whether there are any “prefixed Bounds” of species, any boundaries “made by Nature” (3.6.30:457–8, 3.6.43:466), see Shapiro (1999, 582–7).

³⁹ A parallel claim hold for ideas of modes, such as the ideas of a *triangle* and of *injustice*.

intentionality. In Descartes's case, rather, the divorce is driven by his approach to question (2). This is where he, in turn, advocates a rationalist thesis.

Cartesian rationalism Reflection on our ideas can yield an understanding of fundamental natures, of what makes something what it is.⁴⁰

Yet, quite reasonably, Descartes concedes that the natures he takes to be revealed through rational reflection on our ideas may have little to do with the objects we are currently thinking of, in the *propositional* sense. Rational reflection on the idea the pagan priest expresses by 'Jupiter' can yield an understanding of the nature of the true god, not of the chimerical non-thing his idea is propositionally of. Likewise, reflection on the confused sensory idea the non-philosopher expresses by 'cold' can reveal the nature of a sensation, not the physical quality this idea may be propositionally of. And no amount of reflection on the idea the peasant expresses by 'sun' can reveal the nature of the sun.

The fact that Locke's and Descartes's bifurcations of intentionality are driven by different rationalist commitments explains an important difference between them concerning the ways in which it can make sense to seek an *improved idea of the same object*. According to Descartes, one idea can be better than another in that it more perspicuously reveals a nature. Given that some confused ideas can have two different objects, this results in two very different kinds of cases in which two ideas count as better and worse ideas of the same object. On the one hand, there are cases where the better idea *represents* the same object as the worse idea, even though it counts as the idea of a new object in the *propositional* sense. Such a case is displayed in Table 2.3. Once our pagan makes his idea clear and distinct, it no longer counts, in the propositional sense, as an idea of Jupiter. On the other hand, there are also cases where the better idea remains an idea of the same object in the *propositional sense*, even though it may *represent* a new object. In order to go from believing that the sun is smaller than the earth to understanding that it is "several times larger than the earth" (AT 7:39, CSM 2:27), I must come to employ a radically new "idea of the sun" (in the propositional sense). Above, I suggested that these two ideas have distinct *representational* objects. In any case, it is evident that for Descartes the ends of inquiry can be furthered by replacing a confused idea with a clear and distinct one that remains (in the propositional sense) an idea of the same object.

That is not so for Locke. By his lights, a better idea *can't be* one that more perspicuously reveals an explanatory nature. In this regard, all substance ideas are equally deficient—while all simple ideas and ideas of modes are equally flawless.⁴¹ According to Lockean rationalism, the idea of each substance species

⁴⁰ Locke agrees with this claim in the case of simple ideas and ideas of modes. Here the idea serves not just as nominal essence of a species, but also as its real essence (3.3.18:418, 3.4.3:421, 3.5.14:436). To learn what *makes an action an injustice* or *makes a body crimson*, it suffices to inspect the respective ideas (provided, in the latter case, that one has learned the general lessons of *Essay* 2.8 about color-ideas). Of course, such inspection is a far cry from the serious intellectual work Descartes envisions.

⁴¹ See the previous note.

(in the propositional sense) does reveal informative *membership conditions* for that species, but that is all that can be hoped for. Since he realizes that Cartesians demand more of their clear and distinct ideas, Locke eventually decides it is less misleading to replace the desideratum of “clear and distinct” ideas throughout the *Essay*’s fourth edition with the desideratum of “determinate” ideas, ideas that are consistently expressed by the same name (Epistle:13; cf. 3.11.9:513). At the same time, in a passage quoted in Sect. 2.3.3 above, he inserts a qualification to the effect that “specific *Ideas* of Substances” are subject to an additional norm of perfection (4.12.14:648). What makes one substance idea better than another, assuming both are determinate, is just that the former better supports inductive inferences, as a consequence of conforming more perfectly by correlation to the species it *represents*. While successful inquiry calls for further perfecting the idea to make it more adequately represent that species, Locke would see no point in aiming to formulate a new complex idea that remains an idea of the *same species in the propositional sense*.⁴²

2.5 A Lesson for Current Debates?

What if, together with many philosophers of an “externalist” bent, we reject both Cartesian rationalism and Lockean rationalism?⁴³ I think we can still draw the following lesson from the way Descartes and Locke bifurcate intentionality. Their shared insight is this: the purposes for which we seek an answer to question (1) differ so significantly from the purposes for which we seek an answer to question (2) that we have little reason to expect any neat connection between the kinds of intentionality each answer will invoke. This insight carries consequences for two broad approaches to mental and linguistic intentionality, one “anti-representationalist” and the other “representationalist.”

On the one hand, I have in mind Robert Brandom’s anti-representationalism concerning what I have called *propositional* intentionality, which corresponds to what he calls “discursive intentionality.”⁴⁴ Even if Lockean rationalism is misguided on externalist grounds, Locke is surely right to stress the connection between our ascriptions of *propositional attitudes* and of *truth or falsity* and our interest in clarifying and resolving apparent disagreements. A sophisticated elaboration of this

⁴² In conceding that Locke could even *make sense* of distinct complex ideas that share an object in the propositional sense, I am imagining that he would allow a qualification to (a) of Sect. 2.3.2. Consider the rare case in which distinct substance ideas mutually agree with respect to “necessary connexion,” whence the same objects must conform to each idea. In such a case, I presume, he would count both ideas as being *of the same species* in the propositional sense.

⁴³ For a defense of sophisticated versions of *both* Lockean and Cartesian rationalism, based on an entirely different bifurcation of content, see Chalmers (2012).

⁴⁴ See e.g. Brandom (2008, ch. 6).

theme can be found in Brandom's view of our *propositional* intentional vocabulary as serving to make explicit what is implicit in our practice of challenging or deferring to one another's assertions.⁴⁵ Though Brandom doesn't explicitly motivate his anti-representationalism this way, one might argue that once we appreciate this distinctive function of propositional intentional vocabulary, we will no longer be tempted by "representationalist" explanations what it is to speak or think *that thus-and-so*, and hence to speak or think "of something" in the propositional sense. In response to this motivation for anti-representationalism about Brandomian discursive intentionality, however, a representationalist could concede the point, but insist that she is pursuing question (2), not question (1). That is to say, she doesn't intend her accounts of mental or linguistic representation to serve in explanations of what it consists in to think or speak "of something" in the propositional sense. Furthermore, she can argue that theorists who concentrate exclusively on *propositional* ofness risk neglecting an important, and perhaps largely autonomous, dimension of the bearing of thought and language on the world.⁴⁶

On the other hand, there may also be a lesson for those representationalist theorists whose goal, as articulated by Ruth Millikan, is to "explain the mechanisms, including the contributions of supporting environmental structures, that together account for cases of *proper* cognitive functioning."⁴⁷ One such approach posits mechanisms which, on the occasions they have earned their keep by contributing to successful cognitive functioning, have done so by exploiting *systematically specifiable correlations* between mental or linguistic structures and objects, properties, or states of affairs. This is the broad approach taken by Millikan herself and by Boyd.⁴⁸ Here I would like to call attention to an assumption that is invariably made by such theorists. They assume that the *representational* notions they employ, in addressing question (2), can simultaneously be used to explain those semantic properties that we ordinarily attribute in virtue of our use of the *propositional* notions used in question (1). Such properties include *reference* and *truth conditions* as ordinarily understood: if a mental or linguistic item expresses the proposition that water is liquid, it follows immediately that it has the *truth condition* that water is wet, and that it involves constituents that *refer or apply* to water and to liquid things. The unargued assumption that the representationalist's intentional notions can be used to explain reference and truth conditions may place an unwarranted restriction on the theories of representation at issue.

⁴⁵ See especially Brandom (1994, ch. 8).

⁴⁶ Brandom has recently expressed openness to notions of "representation" that are distinct from the "*discursive* representational" notions that fall directly out of propositional intentionality (2011, 209–19). However, he doesn't raise the question whether the *same* conceptually contentful item might "represent" *distinct objects* in the discursive and non-discursive senses.

⁴⁷ Millikan (1993, 363).

⁴⁸ See also Mark Wilson's case studies in the history of science (2000, 2006). Wilson advocates "the 'correlational point of view': we objectively study how an unfolding reasoning process manages to arrange itself with respect to an independent reality.... This supplies a primitive notion of "truth condition" that is *entirely erected upon the correspondences* uncovered..." (2000, 384).

Understanding how Descartes and Locke are driven to bifurcate intentionality can thus help bring into focus questions that diverse present-day approaches to intentionality have yet to sufficiently address.⁴⁹

References

- Adam, Charles, and Paul Tannery (eds.). 1964–1976. *Oeuvres de Descartes*. Paris: J. Vrin.
- Alanen, Lilli. 2003. *Descartes's concept of mind*. Cambridge, MA: Harvard University Press.
- Anstey, Peter. 2011. *Locke and natural philosophy*. Oxford: Oxford University Press.
- Atherton, Margaret. 2007. Locke on essences and classification. In *The Cambridge companion to Locke's 'Essay concerning human understanding'*, ed. Lex Newman. Cambridge: Cambridge University Press.
- Ayers, Michael. 1991. *Locke: Epistemology and ontology*. London: Routledge.
- Ayers, Michael. 1998. Ideas and objective being. In *The Cambridge history of seventeenth-century philosophy*, ed. D. Garber and M. Ayers. Cambridge: Cambridge University Press.
- Boyd, Richard. 1991. Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies* 61: 127–148.
- Boyd, Richard. 1999. Kinds as the 'workmanship of men': Realism, constructivism, and natural kinds. In *Rationalität, Realismus, Revision*, ed. J. Nida-Rümelin. Berlin: De Gruyter.
- Boyd, Richard. 2010. Realism, natural kinds, and philosophical methods. In *The semantics and metaphysics of natural kinds*, ed. H. Beebe and N. Sabbarton-Leary. New York: Routledge.
- Brandom, Robert. 1994. *Making it explicit: Reasoning, representing and discursive commitment*. Cambridge, MA: Harvard University Press.
- Brandom, Robert. 2008. *Between saying and doing: Towards an analytic pragmatism*. Oxford: Oxford University Press.
- Brandom, Robert. 2011. Pragmatism, expressivism, and anti-representationalism: Local and global possibilities. In his *Perspectives on pragmatism: Classical, recent and contemporary*. Cambridge, MA: Harvard University Press.
- Brown, Deborah. 2007. Objective being in Descartes: That which we know or that by which we know. In *Representation and objects of thought in medieval philosophy*, ed. H. Lagerlund. Aldershot: Ashgate.
- Carriero, John. 2009. *Between two worlds: A reading of Descartes's Meditations*. Princeton: Princeton University Press.
- Chalmers, David. 2012. *Constructing the world*. Oxford: Oxford University Press.
- Clemenson, David. 2007. *Descartes's theory of ideas*. London: Continuum.
- Conn, Christopher Hughes. 2002. Locke on natural kinds and essential properties. *Journal of Philosophical Research* 27: 475–497.
- Cottingham, John, Robert Stoothoff, Dugald Murdoch, and Anthony Kenny, trans. 1984–1991. *The philosophical writings of Descartes*. Cambridge: Cambridge University Press.
- De Rosa, Raffaella. 2010. *Descartes and the puzzle of sensory representation*. Oxford: Oxford University Press.

⁴⁹ Earlier versions of this paper were presented at the Australian National University and the University of New South Wales. For helpful discussion, I would like to thank those audiences as well as the organizers and participants of the Berlin workshop that resulted in this volume, especially Martin Lenz, Antonia LoLordo, Dominik Perler and Anik Waldow. I dedicate the paper to the memory of Joseph L. Camp, Jr., who first made me want to learn from Early Modern theories of intentionality.

- Doyle, John, Trans. 2001. *The Conimbricenses: Some questions on signs* (includes translation of Book 1, chapter 1 of the Coimbran *Dialectic* of 1607). Milwaukee: Marquette University Press.
- Fodor, Jerry. 1998. *Concepts: Where cognitive science went wrong*. Oxford: Clarendon.
- Grüne, Stefanie. 2010. René Descartes. In *Ideen: Repräsentationalismus in der frühen Neuzeit*, vol. 2, ed. Dominik Perler and Johannes Haag. Berlin: De Gruyter.
- Hoffman, Paul. 1996. Descartes on misrepresentation. *Journal of the History of Philosophy* 34: 357–381.
- Hoffman, Paul. 2002. Direct realism, intentionality, and the objective being of ideas. *Pacific Philosophical Quarterly* 83: 163–179.
- Kornblith, Hilary. 1993. *Inductive inference and its natural ground: An essay in naturalistic epistemology*. Cambridge, MA: MIT Press.
- Lennon, Thomas. 2007. Locke on ideas and representation. In *The Cambridge companion to Locke's 'Essay concerning human understanding'*, ed. Lex Newman. Cambridge: Cambridge University Press.
- Lenz, Martin. 2010a. John Locke. In *Ideen: Repräsentationalismus in der frühen Neuzeit*, vol. 2, ed. Dominik Perler and Johannes Haag. Berlin: De Gruyter.
- Lenz, Martin. 2010b. *Locke's Sprachkonzeption*. Berlin: De Gruyter.
- Locke, John. 1975 [1689]. *An essay concerning human understanding*, ed. P. Nidditch. Oxford: Oxford University Press.
- Mattern, Ruth. 1986. Locke on natural kinds as the 'workmanship of the understanding'. *Locke Newsletter* 17: 45–92.
- Millikan, Ruth. 1993. White Queen psychology; or, the last Myth of the Given. In her *White Queen psychology and other essays for Alice*. Cambridge, MA: MIT Press.
- Millikan, Ruth. 2000. *On clear and confused ideas: An essay about substance concepts*. Cambridge: Cambridge University Press.
- Mori, Gianluca. 2012. Hobbes, Descartes, and ideas: A secret debate. *Journal of the History of Philosophy* 50: 197–212.
- Naaman-Zauderer, Noa. 2010. *Descartes' deontological turn: Reason, will, and virtue in the later writings*. Cambridge: Cambridge University Press.
- Nelson, Alan. 1996. The falsity in sensory ideas: Descartes and Arnauld. In *Interpreting Arnauld*, ed. E. Kremer. Toronto: University of Toronto Press.
- Nuchelmans, Gabriel. 1983. *Judgment and proposition: From Descartes to Kant*. Amsterdam: North-Holland.
- Ott, Walter. 2012. What is Locke's theory of representation? *British Journal for the History of Philosophy* 20: 1077–1095.
- Pasnau, Robert. 1997. *Theories of cognition in the later Middle Ages*. Cambridge: Cambridge University Press.
- Pasnau, Robert. 2011. *Metaphysical themes: 1274–1671*. Oxford: Oxford University Press.
- Perler, Dominik. 1996. *Repräsentation bei Descartes*. Frankfurt: Vittorio Klostermann.
- Perler, Dominik. 2010. Einleitung. In *Ideen: Repräsentationalismus in der frühen Neuzeit*, vol. 1, ed. Dominik Perler and Johannes Haag. Berlin: De Gruyter.
- Prinz, Jesse. 2000. The duality of content. *Philosophical Studies* 100: 1–34.
- Shapiro, Lionel. 1999. Toward 'perfect collections of properties': Locke on the constitution of substantial sorts. *Canadian Journal of Philosophy* 29: 551–593.
- Shapiro, Lionel. 2010. Two kinds of intentionality in Locke. *Pacific Philosophical Quarterly* 91: 554–586.
- Shapiro, Lionel. 2012. Objective being and 'ofness' in Descartes. *Philosophy and Phenomenological Research* 84: 378–418.
- Simmons, Alison. Forthcoming. Representation. In *The Cambridge Descartes lexicon*, ed. L. Nolan. Cambridge: Cambridge University Press.
- Stuart, Matthew. 1999. Locke on natural kinds. *History of Philosophy Quarterly* 16: 277–297.
- Suárez, Francisco. 1597. *Disputationes metaphysicae*. In *Disputationes metafísicas*, ed. and trans. S. Rábade Romeo, S. Caballero Sánchez and A. Puigcerver Zanón. Madrid: Gredos, 1960–1966.

- Wee, Cecilia. 2006. *Material falsity and error in Descartes' meditations*. Abingdon: Routledge.
- Wells, Norman. 1984. Material falsity in Descartes, Arnauld, and Suarez. *Journal of the History of Philosophy* 22: 25–50.
- Wells, Norman. 2003. The *Conimbricenses*, Descartes, Arnauld, and the two ideas of the sun. *Modern Schoolman* 81: 27–57.
- Wilson, Margaret. 1990. Descartes on the representationality of sensation. Repr., in her *Ideas and mechanism: Essays on early modern philosophy*. 69–83. Princeton: Princeton University Press, 1999.
- Wilson, Mark. 2000. Inference and correlational truth. In *Circularity, definition and truth*, ed. A. Chapuis and A. Gupta, 371–395. New Delhi: Indian Council of Philosophical Research.
- Wilson, Mark. 2006. *Wandering significance*. Oxford: Oxford University Press.

Chapter 3

Ideas as Thick Beliefs: Spinoza on the Normativity of Ideas

Martin Lenz

3.1 Introduction

In contemporary philosophy, the question whether beliefs are normative is often treated with regard to the fact that beliefs can be true or false. If I believe or say something false, I seem to break a rule or deviate from a standard of semantic correctness. Accordingly, the dispute is about whether there is some sort of social normativity involved here or whether we just happen to deviate from the facts. Normativists tend to argue that the correctness standards are implemented by social practices according to which our linguistic utterances and beliefs are sanctioned; by contrast, naturalists tend to argue that we're just deviating from facts or that the belief producing mechanisms are not properly functioning.¹

One might be unhappy about the state of this discussion for at least two reasons. Firstly, it seems to rely on a controversial gap between nature and normativity, according to which believing is either a normative affair in that it involves social standards or it is something explicable in naturalist terms without any recourse to normative vocabulary. Secondly, it ties the analysis of beliefs too exclusively to truth and falsity, thereby neglecting other factors relevant to the explanation of beliefs. This is why it might be rewarding to look at Spinoza's theory of ideas² for a

Previous versions of this paper have been presented at different workshops in Berlin and Nijmegen. I am particularly grateful to Johannes Haag, Martin Kusch, Dominik Perler, Ursula Renz, Eric Schliesser, Stephan Schmid and Anik Waldow for their insightful comments.

¹ See Engel (2008) for an overview.

² Fruitful interpretations of Spinoza's theory of ideas that point in similar directions are proposed by Della Rocca (2003), Steinberg (2005) and Schmid (2013).

M. Lenz (✉)

Department of the History of Philosophy, University of Groningen,
Oude Boteringestraat 52, 9712GL Groningen, The Netherlands
e-mail: m.lenz@rug.nl

richer, metaphysically grounded and perhaps even more plausible account of the normativity of beliefs.

In portraying ideas as complex propositional attitudes, Spinoza's theory of ideas avoids the shortcomings I mentioned. As we will see, he achieves this by employing a notion of nature which easily integrates normative properties and by depicting beliefs not only as truth evaluable but as the thick states that they probably are: namely as states not only "aiming at truth", but as inherently intertwined with motivating emotions and guiding evaluations. This is why I'd like to call Spinoza's ideas not only beliefs, but – with a nod to Bernard Williams – *thick beliefs*.³

Before looking more closely at Spinoza's account of ideas, I'd like to start the following section with a quick outline of the metaphysical background by reminding you of four basic tenets. In the third part I will distinguish social from natural normativity. Then I move on to Spinoza's theory of ideas. So, the fourth part is about ideas as beliefs, whereas the fifth section aims to show that the content of beliefs is determined by the beliefs' conative elements. In the sixth section I will explain how this way of content determination turns ideas into thick beliefs, before finally discussing whether the normativity of beliefs is a robust or merely a projected feature of thought.

3.2 Four Basic Tenets

Spinoza's theory of ideas should be seen in the context of the main tenets of his *Ethics*. Here Spinoza claims that there is just one substance, namely God or nature. The medium-sized things we commonly deal with – such as stones, cats and humans – are just modes of this substance. This metaphysical thesis is known as Spinoza's *substance monism*.⁴ An important difference between the substance and its modes is that the essence of the substance is said to comprise its existence, it is *causa sui*, whereas finite modes such as humans do not exist as such. Accordingly, the actual essence of these modes is said to consist in their striving (*conatus*) to preserve their being. This is known as the conatus-doctrine.⁵

Now, all things can be considered under at least two attributes, namely thought and extension, such that each mental state, i.e. an idea, can be described as parallel

³ See Williams (1985, ch. 7).

⁴ The following remarks on the main tenets of the *Ethica* are mainly taken from Lenz (2012). References to the *Ethica* (in *Opera* II, ed. Gebhardt 1925) are indicated in the standard way: Roman numerals correspond to parts; abbreviations (often along with Arabic numerals) specify appendix (= app), corollary (= c), definition (= def), demonstration (= d), proposition (= p), and scholium (= s). Translations are taken, sometimes with slight modifications, from Curley (1994). – For the doctrine of substance monism see Spinoza, *Ethica* I p 14–15. See Della Rocca (2008, 46–69), for a thorough exposition of Spinoza's theory of substance and modes.

⁵ See Cook (2006) for a concise exposition of the conatus doctrine.

to a physical state. This tenet known as the *parallelism of mind and body*, implying that the causal explanation of ideas is not to be given with recourse to physical events but to other ideas, just like physical states are causally explained by other physical states. Now, while God or nature is the totality of ideas or parallel physical states respectively, an individual thing such as a stone or a person, is just a part of that totality. You might think of nature as huge net of physical and mental states which are all causally and conceptually related to each other. Your mind, then, is a fairly complex part of this net parallel to the fairly complex state of your body.⁶

This thesis nicely links in with what came to be known as Spinoza's version of *holism*. Our physical and our parallel mental states are embedded in a net of causal or conceptual relations. This means, amongst other things, that the *contents* of our mental states are adequately determinable only in relation to other mental states. Now, since *we* are only part of the whole net, as it were, our cognitions are just as partial and thus (mostly) inadequate. Only from a divine perspective would one be able to access the whole net of causal and conceptual relations. Instead of grasping all the conceptual relations you grasp yourself and the things around you, not from the holistic perspective but in an associative manner.⁷

3.3 Two Kinds of Normativity

Now, where does normativity enter the picture? In the wake of Brandom, Spinoza has been praised repeatedly for his holism of ideational content. Rightly so: as we have just noted, the content of an idea is determinable only with regard to the content of other ideas. At the same time, Spinoza has been chided for his "insufficient appreciation of the *normative character* of the 'order and connection of ideas'".⁸

According to Brandom, conceptual activity is essentially normative in that it rests on *social practices*. In affirming something, for instance, we commit ourselves to something and can be sanctioned by others when violating the rule to which we have committed ourselves. If I affirm that there is an apple on the table but deny that there is any fruit around, then others will set me straight about this.

Now, claiming that conceptual activity is essentially normative not only means that we actually engage in a practice of sanctioning each others beliefs; more importantly, it also means that it is impossible to have beliefs or to be a rational being without such normative practices.

What seems to be missing in Spinoza's account, then, is an answer to the question of what *licenses* the connection of ideas figuring in one's thinking. As we will see in moment, Spinoza *does* have an account of normativity that explains what

⁶ See on the parallelism thesis Della Rocca (2008, 99–104).

⁷ See Spinoza, *Ethica* II p 11 c. On Spinoza's holism see Della Rocca (1996, 68–83).

⁸ Brandom (1994, 93).

grounds this licensing, but one that Brandom would not like all that much, since it is an account, not of *social* but of *natural normativity*.

On my reading, it is the conatus-doctrine, we have just referred to in passing, that lies at the heart of Spinoza's philosophy and of his account of normativity in particular. The term 'conatus' denotes the essence of every individual thing; and this essence lies in the thing's striving for self-preservation and for the increase of its power. Thus, I take it that, for Spinoza, the essence of every thing has basically a teleological structure. Although Spinoza is well-known for his denial of teleological explanations of God or Nature as a whole, he clearly admits that human action is explicable in recourse to this striving.⁹

So we are dealing with *two different kinds of order*: on the one hand, there is the order of nature as a whole, explicable through mechanistic laws; on the other hand, there are individual parts of this nature, such as humans with their individual essences. Accordingly, Spinoza writes in his *Tractatus Politicus*:

So we conclude ... that everyone always, as far as he is in himself, strives to preserve his own existence ... For the bounds of nature are not the laws of human reason, which do but pursue the true interest and preservation of mankind, but other infinite laws, which regard the eternal order of universal nature, whereof man is but a small part.¹⁰

As parts of nature, human beings underlie the mechanistic laws of nature, while their mental and physical behaviour is explicable in terms of their individual law, namely their conatus. Accordingly, nature is basically made up of lots of individuals determined to inter- and counter-act, while striving for their self-preservation.

Taken in themselves, both kinds of laws, that is, the mechanical and individual law of the conatus, are necessary, so neither carries normative force. It is when these laws are looked at in conjunction that normative pressure arises. There is nothing normative about living in a mechanistic universe as such. But if I actually *want to survive* in that universe, then there are things that I *ought or have to* do.¹¹

Here's an easy example: in order to successfully compete in a sports game it is necessary to exercise from time to time. Nothing normative about it; but if I *actually want to win* that game, then I *must* (follow the general law and) make sure to get proper training; otherwise I would violate my commitment to winning the game.

It is at the *intersection* of the order of nature and my individual striving, then, where the normative pressure enters.¹² Thus, there is no special social or even

⁹ The teleological character of Spinoza's conception is of course a matter of highly controversial debate. Yet, it is hard to deny that Spinoza takes recourse to teleological formulations especially with regard to human action; see Viljanen (2011) for an overview.

¹⁰ Spinoza, *Tractatus Politicus* II 8.

¹¹ Accordingly, Spinoza admits to different degrees and thus more or less optimal (i.e. teleologically construed) ways of self-preservation; see Spinoza, *Tractatus Politicus* V.1: "But since the best way of living to assure the utmost self-preservation is that which is framed according to the dictate of reason, it follows that a man or commonwealth acts in the best way, inasmuch as he or it is in the highest degree under his or its own law. For we do not claim that everything of which we say that it is done by right, is also done in the best way. For it is one thing to till a field by right, and another to till it in the best way."

¹² See Spinoza, *Tractatus Politicus* II 8 and Stemmer (2008).

abstract space of normativity. Normative constraints arise in the confrontation of two non-normative orders. Since all of this happens at the natural level, as it were, we might call this *metaphysical or natural normativity*.

Now, if you remember the holistic picture of ideas, you will see quickly how this natural normativity carries over to mental states. With regard to god's mind, all ideas are adequate and well-ordered according to their proper inferential relations, whereas our individual minds are partial, made up of ideas ordered mainly by association. And it is in this context too that normative pressure arises at the intersection of these two orders. The remainder of this paper is devoted to clarifying what this means.¹³

3.4 No Content Without Attitude

Ideas are commonly portrayed as representations. Given the eminent role that ideas play in Spinoza's philosophy, it does not come as a surprise that "representation" is, with regard to Spinoza's theory of mind, taken to be the "essence of the mental."¹⁴ Although this is certainly an apt claim, it does not capture a point that is even more prominent in Spinoza's theory of mind, namely the fact that, for Spinoza, there is no representational content without attitude. Criticising Descartes and his followers, Spinoza writes: "They look on ideas, therefore, as mute pictures on a panel, and . . . do not see that an idea, insofar as it is an idea, involves an affirmation or negation."¹⁵ So, according to Spinoza, our ideas always involve some affirmation or negation. This view closely resembles the contemporary construal of beliefs as propositional attitudes that are analysable into two components, namely content and attitude. According to this view, my belief that there is an apple on the table can be taken to consist of the content *that there is an apple on the table* and my affirmative attitude towards that content, namely my *believing that* things are thus and so. Conversely, my desire that there be an apple on the table consists of the content *that there is an apple on the table* and of my attitude towards that content, namely my *desiring that* things be thus and so.

Now, many contemporary theorists of belief assume that the right order of explanation is to take belief in some sense as *prior* to the bare entertaining of content; that is, the minimal unit of awareness or cognition is not bare content or some part of it, say the concept 'apple', but the whole belief. Brandom notoriously asserts that this claim is a Kantian invention reinforced by Frege, developed in opposition to the pre-Kantian tradition that allegedly started bottom up, that is, with concepts or terms. Brandom sees the privilege that Kant assigned to judgments as a hallmark of the sort of normativity inherent in our conceptual activity: "Kant starts with the judgment because that is the smallest unit for which we can be *responsible*."¹⁶

¹³ The following ideas are partly adopted from Lenz (2012).

¹⁴ See Della Rocca (2008, 90).

¹⁵ Spinoza, *Ethica* II 49s.

¹⁶ Brandom (2001, 80).

Given that Brandom read some Spinoza, it is curious that he ignores Spinoza's claim that our ideas always involve judgments. One reason for disregarding this claim might be that Spinoza does not come up with what Brandom thinks is the right sort of normativity.

As is well-known, however, Spinoza explicitly refutes the Cartesian 'mute picture model'. We cannot *first* just entertain some content and *then choose* to affirm or deny it.¹⁷ In other words: in having ideas we are committed but we are not *freely* committed or taking responsibility. In fact, for Spinoza, the Cartesian mute picture model rests on an illusion produced by a device that is so very dear to contemporary normativists, namely *language*.

It is our thinking in language that deludes us into believing that we can choose to believe what we want and suspend our judgment, or, generally speaking, that our will is free: "Those who confuse words with the idea, or with the very affirmation which the idea involves, think that they can affirm or deny with words something contrary to what they are aware of."¹⁸ A helpful illustration of this confusion is Spinoza's example of a man who states that his courtyard had flown into his neighbour's hen.¹⁹ To be sure, the speaker does not believe what he says. Although the statement is certainly understandable, it does not amount to a belief. The only sense in which it can be affirmed is that it is *uttered*, but an utterance as such does not constitute a mental affirmation.

According to Spinoza then, linguistic expressions and ideas have to be seen as separate items. The critique of the confusion of words and ideas is of course fairly wide-spread already among scholastic philosophers and came up repeatedly, for example, in the debate about the question whether uttering, grasping and affirming content are distinct processes. It seems that there are numerous sentences that we can utter without being able to affirm their content (e.g. "The number of stars is even"). This means that words – unlike the concepts or ideas they are supposed to express – can be recombined without believing in them and sometimes even without being *able* to believe what is said.²⁰ Spinoza not only joins these critics but adds that we cannot have ideas without affirming or negating. This certainly resembles the more modern contention that the "bare entertainment" of content is parasitic on believing.²¹ While it is possible, then, to make utterances regardless of what one actually believes, this is not true of ideas: according to Spinoza, for human beings every idea is a belief; and if human beings believe something, they must believe what they believe. Although the productivity of words might suggest that we can believe at will, we have in fact no choice as to which beliefs we have. So what is it that determines the ideas or beliefs we have?

¹⁷ See Spinoza, *Ethica* II p 43 s.

¹⁸ Spinoza, *Ethica* II p 49s.

¹⁹ See Spinoza, *Ethica* II p 47s.

²⁰ See Lenz (2010, 338 f).

²¹ See, for example, Brandom (1994, 79).

3.5 Content Determination Through Conative Attitudes

As noted above, Spinoza defends a holistic picture of content determination. With regard to the divine intellect at least, the contents of mental states are adequately determinable only in relation to other mental states. But this picture sharply contrasts with the mental lives of human beings. As limited parts of the whole, the contents of our mental states are partial and thus mostly inadequate. But what, then, determines the contents of our ideas?

As we have seen, Spinoza rejects the Cartesian “mute picture model”. Thus, the contents of our beliefs are not taken to be something that can be accessed or entertained independently of our attitudes towards that content. What this suggests is that in order to understand Spinoza’s take on content determination we must look more closely at the nature of propositional attitudes. Although I can utter the string of words “There is an apple on the table” without believing it, I cannot “just think” that there is an apple on the table but at the same time chose or will to believe otherwise. As such, there is no separation between the act of thinking or believing and the will to do so, since, for Spinoza, the “will and the intellect are one and the same;” and they are “nothing apart from the singular volitions and the ideas themselves.”²² This means that ideas taken as beliefs are not merely affirmative acts but tantamount to volitions: to believe is the same as to will. For Spinoza, then, beliefs and acts of will or desire are not different propositional attitudes.

The tenet that beliefs are tantamount to volitions takes us right to the centre of Spinoza’s philosophy of mind and action, namely to his conatus doctrine. Our volitions are basically governed by the conatus. And as we shall see shortly, our beliefs are instantiations of our conatus, our striving. Thus, the answer to the question of what determines the content of our ideas is quite straight-forward: *it is our striving that fixes the content* of our given beliefs.²³

Before we look at this in more detail, I would like to clarify in what sense this explanation of content determination also accounts for the normativity of ideas. As noted at the beginning of this essay, the notion of conatus or striving as such does not render our ideas normative. It is at the intersection of the order of nature and our individual striving that the normative pressure enters. But as we have seen, Spinoza distinguishes between more or less *successful* ways of striving. Given that we are in constant interaction with the world and a mere part of the whole, our striving is also determined by factors external to our conatus, producing inadequate ideas. Since we do not possess divine minds, we cannot see which factors are external and which ones are internal. This means that we do not generally know which factors *actually contribute* to our preservation.

²² Spinoza, *Ethica* II p 49d.

²³ See Spinoza’s variant of the so-called voluntarist principle in *Ethica* III p 9s: “From all this, then, it is clear that we neither strive for, nor will, neither want, nor desire anything because we judge it to be good; on the contrary, we judge something to be good because we strive for it, will it, want it, and desire it.”

Given that we grow up in societies that reinforce conventional associations of ideas, we can be constantly misguided and might fail to form beliefs in accordance with *successful striving*. Believing is thus a *naturally normative* affair in that we can *fail* to believe (and will) what is actually in accordance with our essence as it is located within the order of nature. In other words: if we fail in such ways, we live in accordance with convention. The beliefs we form are not wholly *our* beliefs, as it were, but those reinforced by conventional patterns of associations and the use of language.

Let us now look at these tenets more closely. I have said that to have an idea means to have a belief. To have a belief means to be committed to something (ultimately to one's self-preservation). But let's be careful: of course, Spinoza does not use the term "commitment" in this context. Yet, although Spinoza maintains that we cannot choose to withhold our judgment and thus cannot believe at will, the talk of commitment is quite appropriate, since, as we have just noted, an act of affirmation is not different from an act of will. In Spinoza's own words: "In the mind (by p 48) there is no absolute faculty of willing ... but only singular volitions, namely this and that affirmation ...".²⁴ To have an idea, then, means to will or desire something. But although Spinoza's critique of the Cartesian separation of intellect and will might be justified, it is difficult to see how to make sense of the tenet that the propositional attitudes of believing and desiring should be one and the same. How can, for example, the idea that there is an apple on the table amount to desiring that there be an apple on the table? Does not the desire that there be an apple on the table require that I believe the opposite.

In order to understand the conjunction of desire and belief in Spinoza's account, it is not enough to invoke the conatus doctrine. Rather we must link this doctrine to the parallelism introduced above. According to this doctrine mind and body are parallel to each other. This means not only that for each mental state there is a correlative physical state; it also assigns contents to ideas in the sense that every idea represents the correlating physical state. So according to parallelism, the *essence* of a human mind consists in the idea of its own body: "*The object of the idea constituting the human mind is the body, or a certain mode of extension, and nothing else.*"²⁵ One might say, then, that one's mind is the idea that represents one's body. But since Spinoza rejects the mute picture model this idea is not a mere representation of the body; rather it is the belief or the will that the body exists. Yet, given the holistic thesis according to which we are but one part of the complex causal network of nature, we are constantly interacting with many other external bodies of which our minds also form ideas, namely *inasmuch* as these affect our bodies.²⁶

Now, what does this entail with regard to the thesis that affirmations are tantamount to volitions? Does my belief that there is an apple entail that I want the apple to exist? In some sense this should be the case, but then we seem to end up with the

²⁴ Spinoza, *Ethica* II p 49d.

²⁵ Spinoza, *Ethica* II p 13.

²⁶ See Spinoza, *Ethica* III p 3d.

absurd claim alluded to above, namely that we want the presence of things which we believe to be present at the same time. So how are we to make sense of this claim? The first thing we ought to remember is that we do not grasp the apple as such. As Spinoza makes clear, I can perceive other bodies as existing only insofar as they affect my own body.²⁷ Thus, an idea can be said to indicate *two kinds of content*: primarily the idea indicates my physical state, but inasmuch as this state is caused by external bodies my idea also indicates the external thing. Therefore, my idea of the apple primarily is the idea of an affection of my body.²⁸

With this distinction of primary and secondary indication in place, it is possible to conceive of the identification of affirmations and volitions without running into absurdities. What I primarily want and affirm to exist is my body.²⁹ Yet, along with affirming the existence of my own body I want or affirm the existence of the apple inasmuch as it (positively) affects my body. Generally speaking, then, the *interaction* of my body with other bodies is the *source of the contents* that my mind affirms. Yet again, the question remains: how are these contents determined?

As we shall see, there are at least two ways of describing the process through which these contents are determined. On the one hand, the parallelism suggests, as it were, a basic level at which we can consider the dynamics of content determination arising out of the interactions of bodies. On the other hand, the fact that we consciously evaluate things and ascribe properties to them suggests a refined level of content determination arising from projected norms that we have cultivated through social interactions.

Let's begin our consideration of content determination at the basic or natural level. It will be helpful to start this discussion by looking at Spinoza's famous "physical digression".³⁰ As is well-known, an individual body such as yours can be seen as a highly complex individual that is made up of parts kept in balance according to a rule of motion and rest. This rule can be taken as the conatus under a mechanistic description. Put somewhat anachronistically, one might see this rule (*ratio*) as an individual law governing the physical and biochemical processes, maintaining the body as a functional unit during its interaction with external bodies. This kind of maintenance as a functional unit obviously implies a normative dimension, since some interactions support our bodily functions. Of course, it is impermissibly reductive to ascribe the positive results of such interactions to the external things themselves, but for reasons of simplicity we might say that something external actually contributes and increases the state of our body. An apple a day keeps the doctor away. Other interactions, by contrast, will turn out to decrease the state of our bodies. Now, inasmuch as the motion of bodily parts – no matter whether these parts

²⁷ See Spinoza, *Ethica* II p 26.

²⁸ See Spinoza, *E* II p 16–17 and 26.

²⁹ This does not imply, however, that, in affirming this we have a *conscious* volition or desire. Spinoza makes it clear that this willing concerns our faculty to affirm what is true and not to desire it; see Spinoza, *Ethica* II p 48s, III p 2s and III p 9s.

³⁰ See Spinoza, *Ethica* II p 13–14, and Cook (2006).

are, strictly speaking, parts of my body or external bodies – is in accordance with this rule that helps me to maintain my body, the interaction of parts can be said to belong to the *essence* of the body that is governed by this rule.³¹ So long as this is the case, the striving for self-preservation can be said to be successful. In other words: the success of striving fulfils a natural norm that is instantiated by the given interactions and my will to persist through them.

If we consider this account in the physical digression in conjunction with Spinoza's parallelism, it becomes clear that this rule of motion and rest must have a psychological counterpart. Looking conjointly at the essence of both attributes, body and mind, means of course to focus on the conatus as the essence of the human being.³² A psychological description of the rule and the physical interactions that increase or decrease the integrity of the body would of course have to spell out what the corresponding idea or ideas are. Which ideas are we talking about here? Based on what has been said so far, one answer suggests itself: it is the complex idea of one's own body as existing. Accordingly, Spinoza writes: "Since ... the first thing that constitutes the essence of the mind is the idea of an actually existing body, the first and principal [tendency] of the striving of our mind is to affirm the existence of our body."³³

The mind, then, is a complex idea that strives to affirm the existence of the body. Now it is easy to see in what sense ideas are beliefs or affirmations that are tantamount to volitions. In affirming whatever it is that comes my way I ultimately want my body to exist. If this is correct, then we might say that, according to Spinoza, one's individual beliefs mirror the general teleological structure of one's striving for self-preservation: in affirming p I want q, such that, for instance, in affirming the apple's existence *as it relates to my body*, I want to maintain the existence of my body.

It seems, then, that it is the striving for self-preservation that plays the major role in determining the contents of our ideas. As we have seen, it is primarily the existence of our body that our mind wants; and this idea lies at the heart of all our beliefs and seems to pervade all other contents. In cognitive contact with an apple, for instance, it is the apple's negative or positive impact on my body that makes my conatus pick out those properties of the apple that are relevant to my persistence. Now, while my striving determines the contents of the ideas involved in processes like this, the teleological structure of this striving provides the normative dimension of content. Given my limitations, my ideas can succeed or fail to contribute to the power of my being.

Since the contents of our beliefs are determined by our conatus, our grasp of things seems to focus on what these things provide for us. In this respect, we seem to grasp the things around us in some sense as that what James Gibson introduced as *affordances*, that is, as relational properties disposing us to act. Affordances are,

³¹ See Spinoza, *Ethica* II p 24d.

³² See Spinoza, *Ethica* III p7.

³³ Spinoza, *Ethica* III p 10d.

as Gibson puts it, “what the environment *offers* to the animal, what it *provides* or *furnishes*, for good or ill.”³⁴

Yet, the way in which our interaction with the world has a preserving or decreasing impact on our physical and mental integrity is not the only source of normative constraints to which our ideas respond. As I have pointed out earlier, there is not only a basic or natural normativity that arises at the intersection of our striving and the natural order but also the projected norms that were cultivated by social interaction. After all, as human beings we not only act on affordances but, to some extent, are also conscious of our desires. While this awareness might prompt us to clarify our ideas, it is, according to Spinoza, far more likely that it induces us to build up misguided conventions. This source of normativity will be considered in the following section.

3.6 Conscious Ideas as Thick Beliefs

According to Spinoza, the rise and fall of our power during our interaction with the world is registered by emotions. Given our epistemic shortcomings, we tend to ascribe the positive or negative impact to the things *themselves* and evaluate them accordingly as good or bad. So, when we experience a certain food as increasing our power, we evaluate it as good and tend to associate food that looks or tastes *similar* with the same value. This suggests that superficial similarities do not only play a prominent part in our associations but also in our universal notions.³⁵ Thus, the mechanisms of association and categorization prompt us to evaluate certain things as good or bad, even if they lack any effects or in fact are negatively influencing our well-being.³⁶

Accordingly, we *project evaluative norms* onto things. As social beings with moral and conceptual conventions ingrained in our communicative practices we reinforce and multiply this projective practice in various ways. Thus, the content of our ideas is not only determined by the conatus but, in addition to this, by the emotions and evaluations established by associative patterns and reinforced by our conscious life within society and its customs. In this sense, the conatus driven normativity inherent in our beliefs is blended with a projective normativity. And it is for this reason that I would like to call ideas *thick beliefs*, reminiscent of Bernard Williams’s “thick” ethical concepts: “The way these notions are applied is determined by what the world is like (for instance, by how someone has behaved), and yet, at the same time, their application usually involves a certain valuation of the situation, of persons or actions. Moreover, they usually (though not necessarily directly) provide reasons for action.”³⁷

³⁴ Gibson (1979, 127).

³⁵ See Spinoza, *Ethica* II p 40s1.

³⁶ See Spinoza, *Ethica* III p 15–16.

³⁷ Williams (1985, 129 f).

Since such thick concepts will trickle into our mental lives all the way through our upbringing, we can conclude that previous experiences and hearsay will not only have an impact on our moral judgments but also shape the content of most of our beliefs about the world. Of course, successful striving for self-preservation and striving based on superficial judgments about our well-being can come apart. And according to Spinoza, we are more often guided by superficial similarities than by adequate insights. We are always affected in ways that are good or bad for us, while we might be confused or even ignorant as to *why* we are in fact so affected. But whether or not we are directed at thinking of certain things and properties in the first place will depend what we *take to be* the things' contributions to our well-being.

Spinoza puts this point as follows: "Of ourselves and of a thing we love, we strive to affirm whatever we imagine to affect ourselves or the loved thing with pleasure, and, by contrast, to negate whatever we imagine to affect ourselves or the loved thing with sadness."³⁸

In contrast to the natural normativity rooted in our essence or conatus, this conventionally established and socially consolidated normativity is not based on what actually contributes to our power, but rather on whatever it is we think to be good and right. When our thoughts take this direction, we ascribe the goal-directedness inherent in our striving to the things that, taken as such, have no purpose at all. In striving to be nourished, for instance, we tend to recognize only those features of things that are contributing to this goal, taking the apple not as what it is but as food. Ignoring the response-dependent character of such properties, we take them to be objective features of a detached world around us. Thus, the way we carve up the world and categorise things in the first place hinges on what we take to be good for us, on our acquired preferences and goals. So when we utter the sentence that there is an apple we probably imagine to be stating a fact. But on analysis it turns out that the properties we respond to and the reasons why we have picked out the thing as being thus and so in the first place are determined by our striving for preservation. Even neutral or detached categorisations, such as our sorting things into species and genera, are determined by our striving, then, which is instantiated in thick beliefs that exhibit no principled difference between cognitive, conative and evaluative attitudes.

3.7 Conclusion

As has been argued, Spinoza conceives of our ideas as being inherently normative thick beliefs. The natural normativity of ideas is established by the fact that they can cohere or fail to cohere with our essence as it is located within the order of nature. Conversely, the projected norms inherent in our conscious thick beliefs rest on what we conventionally take to be good for us.

I would now like to conclude this paper by answering a pressing objection regarding the relation between nature and normativity. After all, on the account developed

³⁸ Spinoza, *Ethica* III p 25.

here, nature and normativity seem to be divided at least as it concerns nature, on the one hand, and the kind of projected normativity rooted in convention, on the other. And this seems to suggest that there is an equally neat distinction between two kinds of belief: the conative beliefs, on the one hand, and the emotionally driven and conventionally laden thick beliefs, on the other.

Accordingly, one might think that the normative coinage of conventional thick beliefs is in fact something that we accidentally bring along and that we can rid ourselves of – at least in principle. This would entail that ideas or beliefs are not inherently normative but are so only as a result of our inability to resist the influence that the associative patterns of successive experience exert on us.

Conceived this way, the question is this: what is the status of our normative attitudes? Are they something that would vanish as soon as we clarified the status of our ideas or would they persist in a more robust way? At a first glance, it seems that our normatively-driven categorizations and beliefs are indeed something we make up as long as we keep representing things primarily in the way they affect ourselves. A simple remedy, then, would be to try to attain true beliefs about these things, such that our beliefs are revised in the face of truth.

But such a reply would miss what I have presented to be the very nature of Spinoza's ideas. Since separating one's ideas from one's attitudes seems to require to entertain a content without having an attitude, which would mean to fall back on the Cartesian 'mute picture model' of ideas. What this reveals is that we cannot simply stop forming confused ideas for the simple reason that our (conative) constitution is part of the way we form ideas.

Here, an analogy might help. Can you look at an illustration of the Müller-Lyer illusion without, in some sense, believing that the lines have a different length? Of course, you can counter this belief by invoking different beliefs about the lines, for instance by recalling what has been measured with a ruler. But even so, this belief-revision is made possible through a different route of investigation; it does not rely on the same kind of cognitive processes constitutive of the original belief. Thus, the newly acquired belief does not, strictly speaking, have the same content.

And what is perhaps even more important: in principle, the belief that the lines do have the same length could turn out false, say, for the reason that the measuring instruments have been inadequate. So whether you trust your senses or your ruler will probably depend on what motivates your judgment in the first place. The point, then, is that there is *no independent way* of countering your initial belief.

As we should know by now, Spinoza conceives of ideas in a similar way: an idea's content cannot be entertained without the respective attitude. To rub this point in, Spinoza gives the following example: when looking at the sun, we take it to be, say, 200 ft away. If we get to know the true distance, the error is corrected but we continue to see it as 200 ft away, since we cannot stop seeing it in relation to ourselves. So learning 'the truth' *as such* does not lead to a revision of the belief we have when looking at the sun. Indeed, it might turn out that the initial error was replaced just by another.³⁹

³⁹ See Spinoza, *Ethica* IV p 1s.

We are here faced with conflicting beliefs. And eventually it will be the emotionally and conatively stronger belief (that is, the one having the greater affective force with regard to our striving) that overpowers the other one. So even if our beliefs become more adequate through invoking more and more adequate ideas, they will still be thick beliefs in the sense that they are inherently emotional and evaluative. In other words, it's not truth alone but the actual contribution to our essence (based on the natural normativity) that licenses the connection between ideas.

References

- Brandom, Robert. 1994. *Making it explicit*. Cambridge, MA: Harvard University Press.
- Brandom, Robert. 2001. *Articulating reasons*. Cambridge, MA: Harvard University Press.
- Brandom, Robert. 2002. *Tales of the mighty dead*. Cambridge, MA: Harvard University Press.
- Cook, Thomas. 2006. Der *Conatus*: Dreh- und Angelpunkt der Ethik. In *Baruch de Spinoza – Ethik in geometrischer Ordnung dargestellt*, ed. M. Hampe and R. Schnepf, 151–170. Berlin: Akademie Verlag.
- Della Rocca, Michael. 1996. *Representation and the mind-body problem in Spinoza*. New York/Oxford: Oxford University Press.
- Della Rocca, Michael. 2003. The power of an idea: Spinoza's critique of pure will. *Nous* 37: 200–231.
- Della Rocca, Michael. 2008. *Spinoza*. New York: Routledge.
- Engel, Pascal. 2008. Belief and normativity. *Disputatio* 2(23): 179–202.
- Garrett, Don. 1996. Spinoza's ethical theory. In *The Cambridge companion to Spinoza*, ed. Don Garrett, 267–314. Cambridge: Cambridge University Press.
- Gibson, James. 1979. *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Lenz, Martin. 2010. *Lockes Sprachkonzeption*. Berlin/New York: Walter de Gruyter.
- Lenz, Martin. 2012. Intentionality without objectivity? Spinoza's theory of intentionality. In *Intentionality*, ed. Alessandro Salice, 29–58. München: Philosophia (Basic philosophical concepts).
- Schmid, Stephan. 2013. Spinoza on the unity of will and intellect. In *Partitioning the Soul*, eds. D. Perler and K. Corcilius. Berlin – New York: de Gruyter.
- Spinoza, Baruch de. 1925. *Spinoza opera*, ed. Carl Gebhardt, 5 vols. Heidelberg: Carl Winter [Repr., 1972].
- Spinoza, Baruch de. 1994. *A Spinoza reader*, ed. Edwin Curley. Princeton: Princeton University Press.
- Spinoza, Baruch de. 2002. *Complete Works*. Indianapolis: Hackett. Edition: Michael Morgan (trans: Samuel Shirley).
- Steinberg, Diane. 2005. Belief, affirmation, and the doctrine of *conatus* in Spinoza. *The Southern Journal of Philosophy* XLIII: 147–158.
- Stemmer, Peter. 2008. *Normativität. Eine ontologische Untersuchung*. Berlin/New York: Walter de Gruyter.
- Viljanen, Valtteri. 2011. *Spinoza's geometry of power*. Cambridge: Cambridge University Press.
- Williams, Bernard. 1985. *Ethics and the limits of philosophy*. London/New York: Routledge.

Chapter 4

Three Problems in Locke's Ontology of Substance and Mode

Antonia LoLordo

4.1 Introduction

Locke develops his ontology indirectly, by way of his taxonomy of ideas. The taxonomy of ideas is familiar. All ideas are either simple or complex. All simple ideas are ideas of qualities (2.2.1), and all complex ideas are ideas of substances, ideas of modes, or ideas of relations (2.12.3). Locke's ontology thus consists of four categories:

1. *Qualities*. Examples of qualities are coldness, hardness, and whiteness (2.2.1); perception and volition (2.6.1); and pleasure and pain (2.7.1).
2. *Substances*. Examples of substances are a man, a horse, gold, and water (2.23.1).
3. *Modes*. Examples of modes are beauty and theft (2.12.5); joy, sorrow, hope, and fear (2.20.7–10); a tune (2.18.3); a rainbow (2.18.4); murder (2.12.4); justice (2.30.3); an inch (2.13.4); a triangle (2.12.4); and the number two (2.13.1).¹
4. *Relations*. Examples of relations are father and son, bigger and less, cause and effect (2.25.2); a constable, a dictator (2.28.3); and identity (2.27.1). Locke sometimes treats relations as an independent category and sometimes treats them as a species of modes.² This will be important later.

¹ These are all examples of *mixed* modes. There are also simple modes – “modifications of any one simple idea” (2.13.1) – such as space and extension (2.13.2); place (2.13.7); and hours, days, time, and eternity (2.14.1). I focus on mixed modes, since they cause more problems. For more on simple modes, especially in connection with Locke's views on mathematics, see Carson (2005).

² Cf. Leibniz, *New Essays* 2.12: “This division of the objects of our thoughts into substances, modes and relations is pretty much to my liking. I believe that qualities are just modifications of substances, and that the understanding adds relations ... although relations are the work of the

A. LoLordo (✉)
Department of Philosophy, University of Virginia,
Charlottesville, VA, USA
e-mail: lolordo@virginia.edu

The first category is relatively straightforward. Qualities are a kind of properties. (Roughly, they are the simple, natural, particular properties – either intrinsic or relational – that can be grasped via sensation and reflection.) The difference between items in the first category and items in the second is also relatively straightforward: the contrast between substances and their qualities is, roughly, the contrast between things and their properties.

Matters are more complicated in the case of modes. Locke explains the contrast between substances and modes both in terms of dependence (modes depend on substances for their existence) and in terms of direction of fit (substance ideas are intended to copy features of the world that exist independently of us, while mode ideas are imposed on the world by us). In this paper, I will describe three distinct problems that emerge when you try to explain what modes are and how they relate to substances. First, Locke has no clear explanation of the dependence relation. Second, the direction of fit contrast seems to be in tension with Locke's views on natural kinds: given those views, our ideas of natural kinds cannot have external norms in the way the direction of fit contrast requires. And third, the two contrasts are not equivalent.

I will not suggest a solution to these problems – I cannot find any solution that's compatible with Locke's other commitments – but hopefully the reader will do better. First, however, I will sketch what Locke has to say about modes and their relationship to substances, and then explain why modes are important for Locke and why any problems making sense of their ontology undermine his philosophical system as a whole.

4.2 The Contrast Between Substances and Modes

Locke contrasts substances and modes in two quite different ways. The first – which I will call the dependence contrast – simply falls out of his definition of a mode:

Modes I call such complex ideas, which, however compounded, contain not in them the supposition of subsisting by themselves, but are considered as dependences on, or affections of substances: such as are ideas signified by the words triangle, gratitude, murder, &c. And if in this I use the word mode in a somewhat different sense from its ordinary signification, I beg pardon ... (2.12.4)

Substances subsist by themselves, but modes are dependences on substances.

The second contrast is not articulated in Locke's official definition of a mode, but on my view, it's really the core of the distinction between substances and modes. This is what I think of as the direction of fit contrast.³ Ideas of qualities – simple

understanding they are not baseless and unreal. The primordial understanding is the source of things; and the very reality of all things other than simple substances rests only on the foundation of the perceptions or phenomena of simple substances. Often the same holds with regard to mixed modes, i.e. they ought to be treated rather as relations.”

³ See Bolton (1998) for a discussion of this contrast.

ideas – are supposed to capture the way the world is, independently of us, and Locke insists that they are always successful in so doing (2.30.2, 2.31.2, 2.32.14).⁴ Complex ideas of substances have the same direction of fit. Like simple ideas, they are supposed to capture the way the world really is, independently of us. However, while Locke thinks that we always get it right in the case of ideas of qualities, he thinks we always get it *wrong* in the case of ideas of substances: all substance ideas are inadequate (2.31.6).

Ideas of modes have the opposite direction of fit. We do not pick out something in the world and then design a mode idea to correspond to it. Rather, we design a mode idea to serve certain purposes and then use it to refer to anything out in the world that happens to answer to it⁵:

Nor does the Mind, in these of mixed Modes, as in the complex *Ideas* of Substances, examine them by the real existence of things; or verify them by patterns containing such peculiar compositions in nature. To know whether his idea of adultery or incest be right, will a man seek it anywhere amongst things existing? Or is it true because any one has been witness to such an action? No... (3.5.3).

Instead of first selecting particular existing things and then inventing an idea that matches those things, in the case of mode ideas we make the idea first, and then (perhaps) find existing particulars to match it.

Locke often expresses the direction of fit contrast in terms of archetypes. Both ideas of qualities and ideas of substances are intended to correspond to archetypes in the external world. In contrast, mode ideas have no external, mind-independent archetypes. This is why ideas of modes are always adequate:

Our *complex Ideas of Modes*, being voluntary Collections of simple *Ideas*, which the Mind puts together, without reference to any real Archetypes, or standing Patterns, existing any where, *are*, and cannot but be *adequate Ideas* ... But in our *Ideas of Substances*, it is otherwise. For there desiring to copy Things, as they really do exist; and to represent to our selves that Constitution, on which all their Properties depend, we perceive our *Ideas* attain not that Perfection we intend: We find they still want something, we should be glad were in them; and so are all *inadequate* (2.31.3).

Since there are no mind-independent archetypes that mode ideas can fail to correspond to, they cannot help but be adequate.

Locke sometimes puts this point by saying that mode ideas are always adequate because they act as their *own* archetypes.⁶ For instance, in a passage just preceding the one quoted above, he tells us that ideas of modes are “not ... intended for Copies of Things really existing, but for Archetypes made by the Mind, to rank and denominate

⁴For more on this claim and some of the problems it raises, see LoLordo (2008).

⁵This is consistent with the fact that we're often inspired to design a mode idea because of something we have experienced, as in Locke's example of “the man who first framed the idea of *hypocrisy*,” who “might have ... taken it at first from the observation of one, who made show of good qualities, which he had not” (2.22.2).

⁶Locke also holds that all ideas of qualities correspond to their archetypes, but – since in this case the archetypes are external – he offers quite a different explanation.

Things by” (2.31.2). But either way, what is important for our purposes is that mode ideas have no *mind-independent* archetypes.⁷

Locke also sometimes explains the direction of fit contrast in terms of real and nominal essences. In the case of substances, there is always a distinction between the real essence and the nominal essence:

The measure and boundary of each sort, or species, whereby it is constituted that particular sort, and distinguished from others, is that we call its essence, which is nothing but that abstract idea to which the name is annexed: so that every thing contained in that idea is essential to that sort. This, though it be all the essence of natural substances that we know, or by which we distinguish them into sorts; yet I call it by a peculiar name, the nominal essence, to distinguish it from the real constitution of substances, upon which depends this nominal essence, and all the properties of that sort; which therefore, as has been said, may be called the real essence (3.6.2).

We can also use the language of real and nominal essence when talking about modes. But in the case of modes, the real essence and the nominal essence are always the same:

Thus a Figure including a Space between three Lines, is the real, as well as nominal *Essence* of a Triangle; it being not only the abstract *Idea* to which the general Name is annexed, but the very *Essentia*, or Being, of the thing it self, that Foundation from which all its Properties flow, and to which they are all inseparably annexed (3.3.18).

The identity of the real and nominal essences of modes – like their lack of an external archetype – helps explain why mode ideas are always adequate.

The direction of fit contrast is interesting and important because of the way Locke uses the category of modes, – in particular, the way he uses the fact that mode ideas are always adequate. I gave a number of examples of modes above, including the example of murder and a triangle. And for Locke, the vast majority of moral and mathematical ideas are mixed mode ideas.

This is epistemically important, on Locke’s view. He holds that there is a demonstrative science of mathematics. He also holds that there is, or at least could be, a demonstrative science of morality. These demonstrative sciences are possible because the ideas they involve are ideas that we can know are adequate – ideas of things whose real and nominal essences must coincide. And it is just because moral and mathematical ideas are ideas of mixed modes that we can know they are adequate:

Mixed Modes, especially those belonging to Morality, being most of them such Combinations of *Ideas*, as the Mind puts together of its own choice; and whereof there are not always standing Patterns to be found existing ... may be perfectly and exactly *defined* ... the precise signification of the names of mixed Modes, or which is all one, the real Essence of each Species, is to be known ... Upon this ground it is, that I am bold to think, that *Morality is capable of Demonstration*, as well as Mathematicks: Since the precise real

⁷ As Martin Lenz points out, this marks a clear change from the view expressed in section 26 of Draft A, where Locke refers to moral “Notions or Standards of our actions being not of our own making but depending upon something without us”.

Essence of the Things moral Words stand for, may be perfectly known; and so the Congruity, or Incongruity of the Things themselves, be certainly discovered, in which consists perfect Knowledge (3.11.16)

The claim, then, is that mathematical and moral knowledge is possible because mathematical and moral ideas have the mode direction of fit.⁸

Because of this, Locke's distinction between mode ideas and substance ideas corresponds to a distinction between two different domains of knowledge, broadly construed: on the one hand, mathematics and ethics; on the other hand, natural philosophy. This distinction is central to Locke's philosophical project. Thus, the ontology of Lockean modes is important. And thus, the following three problems with Locke's substance-mode ontology are worrying.

4.3 The First Problem

Here is the first problem: Locke has no clear account of *how* modes depend on substances. It's natural to think that Lockean modes depend on substances roughly as properties depend on the things whose properties they are. But there are four reasons why this cannot be right. I'm not sure that any one of those reasons is compelling all by itself, but together, I think, they are genuinely conclusive.

The first is simple:

- (a) Many of the examples of modes that Locke provides are not things that can be plausibly counted as properties.

A quick look at Locke's numerous examples of modes is enough to establish this. A tune, a rainbow, the number two, and a triangle, for instance, are not properties.

Now, one might object that Locke is just putting his point loosely. What he really means is not that a triangle is a mode, but that the property of being triangular is a mode.⁹ I reply that this would require that Locke is being extremely careless indeed. He cannot just be making an isolated slip in the triangle case, because lots of his examples of modes are things rather than properties. And he tells us explicitly that some mode terms – like some relation terms – are substantives:

[O]ur *simple Ideas have all abstract, as well as concrete Names*: The one whereof is ... a Substantive, the other an Adjective; as Whiteness, White; Sweetness, Sweet. The like also holds in our *Ideas of Modes and Relations*; as Justice, Just; Equality, Equal; only with this difference, That some of the concrete Names of Relations, amongst Men chiefly, are Substantives; as *Patemitas, Pater*; whereof it were easy to render a Reason (3.8.2).

⁸For more on this point and its significance, see LoLordo (2012, Introduction).

⁹I owe this point to an anonymous referee.

Although Locke does not go on to give the reason in this passage, elsewhere he explains why relation terms can be substantives:

[T]he *Ideas* of Relation, may be the same in Men, who have far different *Ideas* of the Things that are related, or that are thus compared. *V.g.* those who have far different *Ideas* of a *Man*, may yet agree in the notion of a *Father*: which is a notion superinduced to the Substance, or Man, and refers only to an act of that thing called Man; whereby he contributed to the Generation of one of his own kind, let Man be what it will (2.25.4).

This makes it hard to believe that when Locke says a triangle is a mode what he really means is that *being triangular* is a mode.

Now for a second reason to deny that modes depend on substances as properties depend on the things that have them:

- (b) It is traditional to conceive of modes as properties, and Locke tells us that he is not using the term ‘mode’ in the traditional sense.

For, in the passage we saw earlier in which he defines the term ‘mode’, he apologizes for using this term in a new way: “if in this I use the word mode in a somewhat different sense from its ordinary signification, I beg pardon” (2.12.4).¹⁰

When Locke was writing, perhaps the most well known usage of the term ‘mode’ was the Cartesian one. Now, the way Cartesian modes depend on substances is, roughly, the way properties depend on things. My thoughts are my modes, and they are properties of me. The grey color of my dog’s fur is one of her modes, and having such grey fur is a property of her.¹¹ So, we should suspect that Lockean modes do not depend on substances in this way.

A third reason to deny that modes are simply properties is related.

- (c) If modes are properties, then qualities should count as modes. But Locke presents the categories of mode and quality as two distinct categories at the same ontological level.

The examples of Cartesian modes just given – my thoughts, the greyness of Maggie’s fur – are examples that Locke would count as examples of *qualities*. And if modes are simply properties, then qualities and modes are the same thing, or at least related as genus and species. But this clearly is not Locke’s picture. Something is either a mode or a quality, not both a mode *and* a quality.

The last reason to deny that modes are simply properties derives from the main use that Locke makes of the category of modes:

- (d) If modes are properties, then it is unclear why modes are the subject matter of the demonstrative sciences.

¹⁰ Part of why Locke’s readers have not typically noticed the problems raised by his ontological category of modes is that they have not recognized just *how* different Locke’s usage of the term ‘mode’ is from his predecessors’ usage. One exception is Ayers (1991, 2.91–109), who points out several of the problems I discuss below.

¹¹ See e.g. *Principles of Philosophy* 1.56 and 1.61 (AT 8a.26 and 29–30).

This is not a knock down argument. Locke's claim that mathematics and morality can be demonstrative sciences because mathematical and moral ideas are ideas of modes is often seen as mysterious or simply absurd. Nevertheless, it would be nice if our way of understanding the category of modes would help us make sense of their peculiar importance for mathematics and ethics.

Thus, I conclude, the distinction between substances and modes is not the distinction between things and their properties. The category of modes includes both things and properties. So, then, in what sense are modes "dependences on" (2.12.4) substances? One way to approach this question is by surveying some different kinds of dependence relations, and seeing which ones might apply. Philosophers have itemized a large number of dependence relations, but we can group them into three main categories: logical, causal, and ontological.

The way in which modes depend on substances obviously isn't *logical* dependence. Modes are not propositions and hence they cannot stand in logical relations to anything. (And the same goes for substances.)

The way in which modes depend on substances is obviously not *causal* dependence either. Now, one of the things that makes understanding the dependence contrast difficult is that although Locke gives lots of examples of modes, he never tells us which substances they depend on. Even the passage that defines modes and then lists triangles, gratitude, and murder as examples of modes does not say which substances they depend on. But the fact that a tune, for instance, causally depends on an instrument and a musician does not seem to be what makes it a mode rather than a substance. And in any case, substances causally depend on other substances, so causal dependence does not even seem like it should be relevant.

So, it must be some kind of ontological dependence. But what kind? Do modes depend on substances for their identity? This doesn't seem right. Granted, it fits some examples. Presumably, the substances on which a murder depends are the murderer and the victim. And it seems plausible that the identity of a particular murder depends on the two people involved. (Perhaps – as in the game *Clue* – on the location and the weapon as well.) But this sort of dependence relation does not seem to apply to all of Locke's examples. I am not entirely sure what substance or substances it is that the number two is supposed to depend on, but in any case, the number two does not seem to depend for its identity on anything other than itself. I would guess that the substance or substances on which a rainbow depends are certain rays of light and bits of atmospheric vapor, but the rainbow does not seem to depend for its identity on which particular bits of atmospheric vapor are involved. And so on.

Perhaps, then, modes depend on substances for their existence. This seems to work better. The constable couldn't exist unless the man existed. The rainbow couldn't exist unless the water vapor and the sun existed. I'm not sure what to say about the triangle and the number two, but the mathematical cases are difficult no matter how you understand the dependence claim: it's just not obvious what substance or substances the mode is supposed to depend on in mathematical cases.

Unfortunately, however, it's not very helpful to say that the way modes depend on substances is that they depend on them *for their existence*. The core notion of a substance is the notion of something that exists independently. Thus, anything that is *not* a substance

must, in contrast, depend for their existence on something else. To say that modes depend on substances for their existence, then, is just to say that modes are not substances. It fails to distinguish modes from qualities, since they too cannot exist without substances.¹² Thus it fails to provide necessary and sufficient conditions for being a mode.

At this point, one might wonder whether I am asking too much of the dependence contrast. Perhaps Locke thinks that different modes depend on the relevant substances in different ways – that he has in mind not one but multiple dependence relations. Perhaps, in other words, the way in which the murder is less fundamental than the murderer and his victim is quite different from the way in which the number two is less fundamental than the substance or substances it depends upon. This would make the category of a mode into a sort of ragbag category: a mode is whatever is not a substance or a quality. This would not help us distinguish modes from qualities or relations. It would not help us understand what the relevant substance is in mathematical cases. And it would not help us understand why mode ideas have the peculiar direction of fit they have.

One might also wonder whether Locke just intends dependence as a necessary condition, not a necessary and sufficient condition. This would eliminate the need to find a sense of dependence in which all modes, but no qualities, are dependent on the relevant substance or substances. I don't think this fits the text all that well. As we've seen, Locke's initial presentation of the distinction between substance and mode is just in terms of the dependence contrast. This makes it sound like anything that is a "dependence on" a substance is a mode. However, I will return to the suggestion that dependence is only a necessary condition in Sect. 4.5 below.

4.4 The Second Problem

The second problem about the distinction between substances and modes concerns the direction of fit contrast. Locke says that substance ideas are intended to correspond to an archetype existing outside the mind, and mode ideas are not. But what exactly *is* the archetype outside the mind? Is he talking about the archetype for the idea of a *particular* substance like the idea of Maggie the dog, or the archetype for abstract ideas like the idea of a dog in general?

The point need not be put in terms of archetypes, although it is convenient to do so. The issue is that it is not obvious whether, in contrasting ideas of substances with ideas of modes, Locke is distinguishing ideas of modes from *ideas of natural kinds*, or simply distinguishing them from ideas of particular substances. It would simplify matters if Locke were just talking about ideas of particular substances. However, this cannot be right, for three reasons. The first is textual:

- (a) In a number of passages, Locke makes clear that he is talking about ideas of kinds of substances, not ideas of particular substances.

¹²One might also worry that it turns composite substances into modes, since composites depend for their existence on their parts. However, Locke makes it clear that the paradigmatic composite substances – things like horses and gold – do not depend for their existence on any *particular* parts. This is the moral of the early sections of 2.27.

Consider passages like this:

Ideas ... of Substances ... have in the Mind a double reference: 1: Sometimes they are referred to a supposed real Essence of each Species of Things. 2. Sometimes they are only design'd to be Pictures and Representations in the Mind, of Things that do exist, by Ideas of those qualities that are discoverable in them. In both which ways, these Copies of those Originals, and Archetypes, are imperfect and inadequate (2.31.6).

Or this:

This piece of Matter, thus denominat'd *Zahab* by *Adam*, being quite different from any he had seen before, no Body, I think, will deny to be a distinct Species, and to have its peculiar Essence; and that the Name *Zahab* is the mark of the Species, and a Same belonging to all Things partaking in that Essence. But here it is plain, the Essence, *Adam* made the name *Zahab* stand for, was nothing but a Body hard, shining, yellow, and very heavy (3.6.47).

The references to species in these passages make clear that Locke is talking about the idea of the kind, not just the idea of particulars.

A second reason that Locke must be contrasting two kinds of abstract or general ideas is that he alternately expresses the direction of fit contrast in terms of archetypes and in terms of real and nominal essences. Only substance ideas have mind-independent, external archetypes to which they are supposed to correspond, and only substances ideas are of things with real essences distinct from their nominal essences:

- (b) Since Locke articulates the direction of fit contrast in terms of real essences and holds that individuals only have real essences relative to a kind, he must be talking about the archetypes of general ideas of substances.

That individuals only have real essences when considered in relation to a kind is the moral of a well-known passage:

'Tis necessary for me to be as I am, GOD and Nature has made me so: But there is nothing I have, is essential to me. An accident, or Disease, may very much alter my Colour, or Shape; a Fever, or Fall, may take away my Reason, or Memory, or both; and an Apoplexy leave neither Sense, nor Understanding, no nor Life. Other Creatures of my shape, may be made with more, and better, or fewer, and worse Faculties than I have: and others may have Reason, and Sense, in a shape and body very different from mine. None of these are essential to the one, or the other, or to any Individual whatsoever, till the Mind refers it to some Sort or *Species* of things ... So that if it be asked, whether it be *essential* to me, or any other particular corporeal Being to have Reason? I say no ... (3.6.4).

Thus, again, it seems that Locke's intention must be to contrast ideas of modes with ideas of natural kinds.

A final reason that Locke must be thinking about substance kinds, not particular substances, is this:

- (c) Particular *mode* ideas are meant to correspond to archetypes outside the mind too: we lose the direction of fit contrast between substances and modes if we restrict ourselves to ideas of particulars.

My idea of murder or adultery in general is not intended to capture some independently existing natural kind. But my idea of a *particular* murder – say, the assassination of JFK – is intended to capture what that event was really like.

For these three reasons, it seems implausible that in contrasting substances and modes Locke is contrasting two types of particular ideas. He must instead intend a contrast between two different types of abstract ideas: ideas of natural kinds like horses and gold, and ideas of kinds that are dependent upon our way of conceptualizing the world for their existence. This is a reasonably intuitive contrast, and one that many philosophers have drawn. But it would be surprising for *Locke* to draw it. For the whole tenor of his discussion of so-called natural kinds is that they too are dependent upon our way of conceiving the world. Species “are the Inventions and Creatures of the Understanding, made by it for its own use” (3.3.11).

We need to be a bit careful here. Locke does not think that we typically realize that species are the workmanship of the understanding. He thinks that what we usually believe is that there are sharply delineated natural kinds out in the world, and that our natural kind terms correspond to them. In other words, we typically intend for our nominal essences – our abstract ideas – to correspond to real essences. Thus in our ordinary use of natural kind terms, we *are* “desiring to copy Things, as they really do exist.” We are trying to capture some “real Archetypes, or standing Patterns” existing outside the mind. Of course, Locke thinks that we always fail to do so, but that doesn’t matter for present purposes. Remember, Locke draws the direction of fit contrast at the same time that he tells us that ideas of substances are always inadequate and ideas of modes are always adequate. The problem comes in when we think about how Locke himself uses – and recommends using – substance terms.

Now, there is some disagreement about whether Locke’s remarks about species boundaries are intended as metaphysics or epistemology. I read him as expressing skepticism about the existence of genuine natural kinds, but on some interpretations, he is merely pointing out that we lack knowledge of genuine natural kinds.¹³ But there is a problem either way. If Locke thinks there are no sharply delineated natural kinds, no archetypes in nature, then he is obviously not using natural kind terms in an attempt to capture some mind-independent archetype. If, on the other hand, he thinks that there may be natural kinds but denies that he has any knowledge of them, it would still make no sense for him to be using natural kind terms in an attempt to capture some mind-independent archetypes. Given what Locke thinks about natural kinds and our knowledge of them, he himself ought to be using natural kind terms with the mode direction of fit, not the substance direction of fit.

I asked whether the contrast between substances and modes was a contrast between two types of particulars, or a contrast between two types of kinds – natural and conventional. I argued that there are problems making sense of the direction of fit contrast whichever way you go. Here, it’s also worth noting that the distinction between abstract ideas of substance kinds and ideas of particular substances is not always clear. When we are talking about substances designated by mass nouns rather than count nouns, the contrast cannot be made. And Locke is quite fond of using mass kinds as examples: gold and water are the most common. Indeed, the prevalence of such examples is part of why the text leaves it ambiguous whether Locke is talking about the archetypes of individuals or the archetypes of natural kinds.

¹³ For a recent and well-articulated epistemic interpretation, see Anstey (2011, chapter 11).

4.5 The Third Problem

The third problem is one many readers will already be familiar with, since Leibniz points out one version of it in the *New Essays*. Although Locke presents the dependence contrast and the direction of fit contrast as two different explanations of the distinction between substance and mode, the two contrasts do not actually carve up the world in the same way. Certain examples with the mode direction of fit do not look like dependences in any plausible sense, and certain paradigm dependences have the substance direction of fit.

Leibniz gives the example of a dependence with the substance direction of fit:

Illnesses imitate substances, so to speak, in such a way that an illness resembles a plant or animal which requires an account all its own. That is, illnesses are 'modes' or ways of being which fit what we have said about bodies or substantial things, a quartan fever [malaria] being as hard to understand thoroughly as is gold or quicksilver (Leibniz, *New Essays* 4.7/426).

The idea of malaria seems to have the substance direction of fit: we design it to fit a phenomenon that we have observed in the external world. Malaria has a real essence that is distinct from its nominal essence, one that is open to empirical investigation. And it is clearly possible to have an inadequate idea of malaria.

Leibniz goes on to challenge Locke's conception of the distinction between substances and modes in general:

The patterns of one of these kinds of idea are just as real as the patterns of the other. The mind's qualities are no less real than the body's. True, one does not see justice as one sees a horse, but one understands it as well, or rather one understands it better. Whether or not one gives thought to it, justice inheres in actions as much as straightness and crookedness do in motions. To show you that my opinion is shared by others, even the ablest and most experienced in human affairs, I need only appeal to the authority of the Roman jurists, who have been followed by all the others. They speak of these 'mixed modes' or 'moral beings' of yours as *things*, specifically *incorporeal things*. For example, they speak of legal rights, such as a right of way over a neighbor's land, as incorporeal things which can be owned, can be acquired through long use, can be possessed, and can be claimed by legal action (Leibniz, *New Essays* 3.5/303).

Unfortunately, Leibniz's challenge rests on a misunderstanding. For one thing, it is simply false that, as Leibniz assumes at the beginning of the passage, modes are mental entities while qualities are bodily entities.¹⁴ (This misunderstanding may derive from the fact that Leibniz is focusing on one particular sort of mixed modes, moral entities, in this passage.) For another, Leibniz misunderstands the distinction between substance and mode when he assumes that a thing cannot be a mode. For Locke would agree both that something like a property right is a mode, and that it is a thing in the sense that it can be possessed.

Nevertheless, the disease example is powerful. The fact that Leibniz uses a medical example is interesting, given Locke's medical background. I am unsure whether

¹⁴It is false that all modes are mental for Locke: consider the rainbow. And it is also false that all qualities are bodily: consider perceiving and reflecting, or pleasure and pain.

Locke and Leibniz would think of a disease as a set of symptoms, the underlying cause of those symptoms, or something else entirely.¹⁵ I will not dwell on this, though, since once you see Leibniz's example it is easy enough to think of other, non-medical examples.

Consider, for instance, a hurricane, or an earthquake. Hurricanes and earthquakes are clearly dependences, in the relevant sense. The earthquake depends for its existence on the earth that moves; the malaria depends for its existence on the sick person.¹⁶ However, hurricanes, malaria, and earthquakes seem to have the substance direction of fit. They have real essences distinct from their nominal essences that can be investigated empirically. Indeed, in terms of direction of fit they seem to be as clear cases of substances as paradigmatic examples like horses and gold. Thus, Leibniz's example and others like it seem to show that Locke's two different ways of conceiving the distinction between substance and mode are actually ways of conceiving *two different distinctions*.

Notice that the disease and earthquake examples cause problems whether you think of dependence as a necessary and sufficient condition for being a mode, or merely as a necessary condition. For even if something can be dependent in the relevant way without being a mode – even if there's no significant difference between the way that modes and qualities depend on substances – it's clear that for Locke, *in*-dependence is necessary for being a substance. Thus, hurricanes, malaria, and earthquakes cannot be substances. But they cannot be modes either, because they lack the mode direction of fit. Thus they are uncategorizable. But clearly, Locke thinks his contrast between substances, qualities, modes, and relations is exhaustive: else his taxonomy of ideas would be incomplete.

Just as there are examples of paradigmatic dependences with the substance direction of fit, there are examples of paradigmatic independences with the mode direction of fit. One example of an independent thing whose idea has the mode direction of fit is a watch. Locke introduces the example of a watch for quite a different purpose in the discussion of individuation in 2.27. There, he assimilates watches to animals, thus apparently treating them as substances:

[W]hat is a watch? It is plain 'tis nothing but a fit organization, or construction of parts, to a certain end, which when a sufficient force is added to it, it is capable to attain. If we would suppose this machine one continued body, all whose organized parts were repaired, increased, or diminished, by a constant addition or separation of insensible parts, with one common life, we should have something very much like the body of an animal ... (2.27.5)

¹⁵ The term 'quartan fever' – meaning a fever that recurs on the fourth day – suggests that the disease was originally conceived as the set of symptoms. But Leibniz's view that malaria has a real essence distinct from its nominal essence suggests that he conceives of malaria as the underlying cause of those symptoms instead.

¹⁶ If you think that the malaria is the underlying parasite, then you might deny that the malaria depends for its existence on the sick person. (You would thereby turn malaria into a substance.) But the same move cannot plausibly be made for hurricanes and earthquakes.

In terms of the contrast between dependent and independent things, a watch is clearly a substance.¹⁷ But consider the abstract idea of a watch. The idea of a watch is something that is made by us for our own purposes, rather than copied from nature. The idea of a watch must be adequate: if there are no things in the world that fit the idea, this does not show that the idea is flawed. (It shows that no one has yet succeeded in making a watch.) The archetype to which the idea is supposed to correspond is not some mind-independent natural kind but something within our own minds: watches aren't natural kinds. Thus, the abstract idea of a watch – like the idea of *any* artifact – has the mode direction of fit. This shows, again, that Locke's two different ways of characterizing the distinction between substances and modes are not equivalent.

4.6 Conclusion

I have argued that Locke's distinction between substances and modes is important because of its key role in the epistemology of mathematics and ethics. I have also described three problems with that distinction. The first is a problem with the dependence contrast between substances and modes. Locke cannot clearly articulate the sense in which modes depend on substances: all he has to offer is the vague and unhelpful characterization of it as a sort of existential dependence. The second is a problem with the direction of fit contrast between substances and modes. He spells out that contrast by explaining that substance ideas are intended to correspond to external, mind-independent archetypes. It seems that Locke must intend to contrast ideas of substance kinds with ideas of conventional kinds, but this contrast loses its force given his conventionalism about natural kinds. The third problem is that the two contrasts do not actually seem to count the same things as modes and substances. Certain examples are modes according to the dependence contrast and substances according to the direction of fit contrast, or vice versa.

I do not claim that these problems are insoluble. However, at the moment, I see no way to solve them. The direction of fit problem strikes me as particularly intractable. In his discussions of modes and of moral and mathematical knowledge, Locke insists that substance ideas are underwritten by the world in a way in which mode ideas are not. This is what grounds the special epistemic status of morality, which is more certain than natural philosophy. But in his discussions of natural kinds, he calls into question the claim that substance ideas are underwritten by nature into question – thereby calling into question the direction of fit contrast and the distinct epistemic status of morality. This reflects what I see as Locke's deep ambivalence about the extent to which nature can impose norms on human thought.

¹⁷Of course, watches depend for their existence on the existence of certain material parts. But this is no different from the way animals depend for their existence on the possession of some material body. That's the point of the comparison between animals and watches Locke is making here.

References

- Anstey, Peter. 2011. *John Locke and natural philosophy*. New York: Oxford University Press.
- Ayers, Michael. 1991. *Locke*, vol. 2. London: Routledge.
- Bolton, Martha. 1998. Substances, substrata and names of substances in Locke's *essay*. In *Locke*, ed. Vere Chappell. New York: Oxford University Press.
- Carson, Emily. 2005. Locke on simple and mixed modes. *Locke Studies* 5: 19–38.
- LoLordo, Antonia. 2008. Locke's problem concerning perception. *Philosophy and Phenomenological Research* 72.3: 705–724.
- LoLordo, Antonia. 2012. *Locke's moral man*. Oxford: Oxford University Press.

Chapter 5

Kant on Imagination and the Natural Sources of the Conceptual

Johannes Haag

It is well-known that in his theoretical philosophy Kant puts forward an approach to knowledge involving two ‘stems’ that necessarily interact to generate knowledge of a world existing independently of the subject of experience. This comprises, on the one hand, the receptivity of our sensibility, somehow responsible for intuitive representations and, on the other hand, the spontaneity of our understanding, conceived as a faculty of conceptual representations. It seems that the differentiation between these two ‘stems’ aligns quite neatly with the distinction between natural and normative influences on our knowledge respectively.

One of the persisting questions of Kant scholarship, however, is how nature and normativity so conceived can interact in such a way as to produce *empirical* representations that are simultaneously shaped by our conceptual constraints. In other words, while conceptual resources seem to have a significant influence on those representations, at the same time we need to ask how this representational spontaneity can, as Wilfrid Sellars once put it,¹ be ‘guided from without’ by a receptive sensibility in order to guarantee that the ensuing representations are truly *empirical*. Consequently, sensibility and understanding, the receptive and spontaneous faculties, must *interact* if they are to generate conceptual empirical representations, i.e., normativity constrained by nature.

To make this possible, Kant introduces a further faculty in his system that, guided by the understanding, allows subjects of experience to ‘synthesize’ or unify the representational input of sensibility into conceptually shaped representations, namely, the faculty of *imagination*. My paper will be devoted to outlining the philosophical theory behind these ideas. Exegetically, it will focus on the *Critique of Pure Reason (CPR)*, while a broadly Sellarsian interpretation will serve as a systematic background. In this way we can at least begin to do justice to the complex function of the imagination as an intermediary between conceptual norms and nature in Kant’s philosophy.

¹ Cf. Sellars (1968, 16).

J. Haag (✉)

Department of Philosophy, University of Potsdam, Potsdam, Germany
e-mail: jhaag@uni-potsdam.de

5.1 The Faculty of Presentation

What then is the role of the imagination in Kant's system? In its most general form, it is a *faculty of presentation (Darstellung)*.² Kant in the third *Critique* introduces the concept of *presentation* as follows:

If the concept of an object is given, then the business of the power of judgment in using it for cognition consists in presentation (exhibitio), i.e., in placing a corresponding intuition beside the concept. (CPJ 5:192)

That Kant, in the context from which this quotation is taken, ascribes this activity of presentation *not* to the imagination but to the power of judgment may seem rather surprising, especially since in other places in the same text Kant is very clear that presentation is a (central) business of *imagination*. For instance, in a later paragraph of the *Critique of the Power of Judgment* he writes: “[T]he faculty of presentation is the imagination” (CPJ 5:232) (another case in point can be found in the notorious *Deduction of the Analytic of the Aesthetic Power of Judgment*, Cf. CPJ 5:287.) What is furthermore of importance for us is Kant's emphasis on the claim that the faculty of *apprehension* is “one and the same” (CPJ 5:279) with the faculty of *presentation*. Apprehension, however, will in turn be clearly identified as part of the functioning of the imagination. Consequently, for the purposes of the task at hand we are justified in sticking to the *imagination* as the faculty that does the presenting and so to returning to the task of characterizing presentation.

In order to provide a Kantian answer to this question it will be helpful to explicate the meaning of the claim that the faculty of imagination is at the same time the faculty of apprehension. Imagination in Kant's Critical Philosophy is first and foremost the faculty of *synthesis*. Its central epistemological function consists in the synthetic construction of intuitions. It can do this, however, only in accordance with – and with the help of – concepts. In a complex ‘threefold synthesis’ the imagination takes up the sensibly given material into consciousness and restructures it in accordance with the forms of intuition specific to our (human) sensibility as well as in accordance with the categories. This complex process is called the synthesis of *apprehension*.³

In the course of this process the subject synthetically apprehends the sensibly given material and out of it constructs complex representations that are then *taken* as objects:

Here that which lies in the successive apprehension is considered as representation, but the appearance that is given to me, in spite of the fact that it is nothing more than a sum of these representations, is considered as their object, with which my concept, which I draw from the representations of apprehension, is to agree. (A191/B236)

Are those representations Kant's *intuitions*? Not quite. In the interpretation I would like to defend they serve rather as the sensible objects to which our intuitive

² CPJ 5:232. Cf. 28.1:235 ff.

³ At least in the 2nd edition of the *Critique of Pure Reason*. Cf. B162. In the first edition the synthesis of apprehension is only the first part of the threefold synthesis. Cf. A98–100 and Sect. 5.3 below.

representations demonstratively refer. According to this interpretation, which in this respect closely follows Wilfrid Sellars's elaborate interpretation of Kant's conception of intentionality, it is the basic function of intuitions to bring an object before the mind for consideration.⁴ In order to fulfil this task then, intuitions presuppose as it were sensible models construed by the imagination according to recipes of construction provided by the empirical concepts in question – models that are, as Kant writes in the *Schematism*, products of methods for providing concepts with pictures.⁵ Those sensible, representational models (Sellars's *image-models*) – *taken by the representing subject as objects* – thus turn out to be the proper candidates for demonstrative reference. Imagination is thus a faculty that literally provides concepts with pictures made from *sensible* material in accordance with *conceptual* recipes for construction: it is in this sense a faculty of *presentation*.

Let me try to flesh out these rather sketchy remarks thus far in more detail. The exposition will have two parts: firstly, in order to clarify how normativity shapes nature in terms of the specific way in which the imagination synthesizes the material receptively given in the presentation of objects, I will try to elucidate the concept of an image-model and its relationship to intuitive representations in a more systematic manner. At the same time, I will also pay closer attention to the overall picture Sellars gives of this part of the Kantian system. In a second, more explicitly exegetical step I will try to shed further light on the concept of synthesis presupposed in this process, which ultimately provides us with empirical representations of objects of experience.

5.2 Image-Models

The conception outlined above regarding the relationship between image-model and intuition relies heavily on what I take to be Sellars's late, more sophisticated⁶ and thoroughly Kantian conception of perception and, in particular, his concept of an image-model that is introduced in order to make comprehensible the conceptually guided, and at the same time empirically constrained, presenting activity of the Kantian imagination. Image-models are, in the first approximation, complex, three-dimensional images of objects and, as such, are the result of the operation of a conceptually guided imagination on non-conceptual sensory input.⁷ Now, if

⁴Cf. Sellars (1978a, § 48).

⁵A140/B179/80.

⁶Jay Rosenberg and John McDowell, both of whom have worked extensively on Sellars' Kant-interpretation, share this estimation. Cf. Rosenberg (2007b, 240) and McDowell (2009, 114). Rosenberg in this paper gives an excellent sketch of Sellars's theory of perception in general and his conception of image-models as products of productive imagination in particular. The way he relates the latter to Kant's account of the threefold synthesis (*ibid.*, Fn.12) is, however, quite problematic, as we will see below.

⁷Sellars introduces this activity of the imagination by way of "phenomenological reflection" (Sellars 1978a, §§ 3, 27), but the implications of the concepts thus gained will prove paramount and, moreover, they are in line with what Kant has to say about the empirical activity of the productive imagination.

image-models were the result of a conceptual shaping of strictly non-conceptual sensory input via the activity of what Sellars, following Kant, calls⁸ the *productive imagination*, then we should start by asking: what is the nature of the sensory input fed into this shaping? Purely receptive *sensations* (or sense-impressions) are conceived upon first analysis as states of perceiving subjects, which are the *effect* of our senses being affected by external objects. Moreover, these states themselves *contain* the *sensibilia* which the perceiving subject conceives as the properties of the objects of perception. As sensations, they are “non-conceptual states of consciousness... *none* of which are apperceived” (Sellars 1968, 10).⁹

Given this picture of the *material* upon which the productive imagination operates, what happens to it in the synthesizing process? The synthesizing activity of the productive imagination *forms* this receptive sensory input such that it becomes the qualitative content of a spontaneously, and hence conceptually-structured, complex *image* of a three-dimensional object. This object is represented with its sensory properties and is pictured from the perspective of a perceiving subject. Image-models are thus the conscious shapings of the unconscious receptive input that is situated below the line which separates not only receptivity from spontaneity but also sub-conscious mental states from conscious ones.¹⁰

But in what sense is the content of those images conceptually shaped? Image-models themselves are conceptual as well as sensory. The productive imagination could not form image-models unless it did so according to recipes provided by the understanding – recipes that are part of our empirical concepts of objects.¹¹ Those recipes are therefore designed to play a role exactly corresponding to Kant’s concept of an empirical *schema*: “This representation of a general procedure of imagination in providing an image of a concept, I entitle the schema of this concept.” (A140/B179/80)

The schema is thus a conceptual recipe for forming those sense-impressions restricted by the Kantian *mathematical*, though *not* the *dynamical* categories and their corresponding transcendental schemata.¹² Image-models, consequently, are the result of a spontaneous, conceptual shaping of sensations, which in turn is the result of our receptive faculty being affected by things-in-themselves. Consequently, the properties of the image thus construed are only¹³ their *sensibilia*. They comprise,

⁸ Cf. Sellars (1968, 4; 1978a, §§ 1, 28–36).

⁹ We will find that this statement is in need of elaboration with respect to the Kantian approach. Cf. below Sect. 5.5.

¹⁰ For the metaphor of a line that is supposed to separate what is situated ‘above’ it (in the sphere of the conceptual or spontaneous) and that what lies ‘below’ it as *purely* receptive cf. McDowell (2009).

¹¹ Sellars (1978a, §31). Cf. the discussion in McDowell (2009, 114).

¹² Sellars does not make this important difference explicit, but ultimately gives a description of the role of empirical and transcendental schemata that fits this Kantian distinction when he distinguishes “empirical structure” from “‘categorical’ features” (Sellars 1978a, § 39) Cf. *ibid.*, §§ 22, 24. For the distinction in the *Critique of Pure Reason* cf. A160/B199.

¹³ Cf. Sellars (1978a, §22).

however, not only the properties actually perceived, but also the merely imagined sensible properties that we represent the object as having:

We see the cool red apple. We see it as red on the facing side, as red on the opposite side, and as containing a volume of cool white apple flesh. We do not see of the apple its opposite side, or its inside, or its internal whiteness, or its coolness, or its juiciness. But while these features are not seen, they are not merely believed in. These features are present in the object of perception as actualities. They are present by virtue of being imagined. (Sellars 1978a, §21)

Image-models are a blend of features seen *and* features imagined, a “*sensing-cum-imaging* a unified structure” (ibid., §24). They are what we *take to be* the objects of which we are directly aware in sensory or, as Sellars sometimes puts it, perceptual consciousness: “[A]lthough the objects of which we are directly aware in perceptual consciousness are image-models, we are not aware of them as image-models” (ibid., § 27). The mental states, which are the *takings* of those image-models as objects, are what Sellars calls ‘perceptual takings’¹⁴ and what Kant – or at least the Kant of the Sellarsian interpretation introduced above – calls *intuitions*.¹⁵ Hence, intuitions are the *takings* of the image-models *as* objects of experience in a demonstrative thought, and image-models in turn are, again in Kantian terminology, the appearances that are taken to be objects in an intuition of an object of experience.¹⁶

Intuitions are distinct from image-models since more is necessary for taking something to be an object of experience than to merely ascribe to it sensible properties. Intuitions represent their objects with causal and dispositional properties that we “do not see *of* [those objects]... though we see them as having them” (Sellars 1978a, §22), whereas image-models only contain properties of actual or possible sensory experience. Intuitions furthermore are representations of objects of experience whose *esse* essentially is not *percipi*, while image-models are representations of objects whose *esse* is *percipi* in that they are essentially perspectival objects.¹⁷ Image-models are, as it were, objects without objectivity: they always incorporate the perspective of the perceiving subject.¹⁸ The schemata that provide the recipes for their construction are never just schemata of objects but always of objects “in such-and-such relation to a perceiver” (Sellars 1978a, § 34).

Most importantly, intuitions are *not* sensory representations: they are representations that serve to make a conceptually-laden demonstrative reference to an image-model, which is taken to be an object of experience, thereby “bringing a particular object before the mind for its consideration” (Sellars 1978a, § 48). In an intuition we take a complex sensory object to be an object of experience. In this way, an intuition can serve as the subject of a perceptual judgment that guarantees direct

¹⁴ Cf. e.g. Sellars (1976, § 53; 1978a, § 10, 50).

¹⁵ Cf., for instance, Sellars (1976, § 24).

¹⁶ Cf. the quote above from A 191 / B 236.

¹⁷ Cf. ibid. § 28; Sellars (1976, § 51).

¹⁸ Cf. Sellars (1978a, § 28).

contact between the ensuing judgment and what we take to be (part of) the world outside via its demonstrative aspect (though what we take to be part of the world outside is ultimately not in any meaningful sense ‘out there’¹⁹).

The conceptual content of intuitions enables them to figure in Sellars’s theory as perceptual *takings* in the full sense of that term, namely, as having a proto-judgmental form.²⁰ That is why, as Sellars points out, we may think of this kind of taking as believing, although we must think of it as ‘believing in’ rather than ‘believing about’ or ‘believing that’²¹:

What is taken or, if I may so put it, believed in is represented by the complex demonstrative phrase; while that which is believed about the object is represented by the explicitly predicative phrase which follows. Perceptual takings, thus construed, provide the perceiver with perceptual subject-terms for judgments proper. (Sellars 1978a, § 10)

Again, this sharply distinguishes intuitions from image-models, which do not serve as means of reference to objects, but as those (mis-taken²²) objects themselves. As McDowell correctly puts it: “Sellars does not consider claim-containing occurrences that are themselves shapings of sensory consciousness” (McDowell 2009, 122).

Naturally, there are further questions concerning the relationship between image-models and intuitions. One may ask: can the recipe (schema) used in the generation of the image-model be simply identified with the “demonstrative conceptualization” (Rosenberg 2007a, 273) (that is, the intuition), as Jay Rosenberg claims? I am skeptical and would like to suggest instead that this particular use of concepts and their correlated schemata guides the activity of the productive imagination, which generates *both* the construction of image-models and the demonstrative reference to this image-model as an object of experience. Indeed, Sellars claims that the productive imagination “is a unique blend of a capacity to form images in accordance with a recipe, and a capacity to conceive of objects in a way which supplies the relevant recipes” (Sellars 1978a, § 31).

But here the capacity to conceive of objects should not be identified with the capacity to demonstratively refer to objects (although the latter presupposes the former). It should be understood as the capacity to use concepts in general – a capacity that is presupposed not only in the construction of image-models, but equally so in the forming of an intuitive representation of an object. This explains why Sellars himself proceeds by speaking about concepts and schemata without mentioning demonstrative reference:

Kant distinguished between the concept of a dog and the schema of a dog. The former together with the concept of a perceiver capable of changing his relation to his environment implies a family of recipes for constructing image models of perceiver-confronting-dog. (Sellars 1978a, §31)

¹⁹ Cf. Sellars (1968, 48/9; 1963, 97).

²⁰ Cf. Sellars (1978b, 280/1).

²¹ Cf. Sellars (1982, 87).

²² For the conception of mis-taking cf. Sellars (1982, 109).

There is no reference here to the use of concepts in their capacity to intuitively refer to an object, but only to concepts and the related schemata. Both concepts and schemata are simply presupposed in ‘providing concepts with images’ *and* in intuitive reference.

I am prepared to endorse much of what Sellars has to say about the role of imagination in both the construction of image-models and the conceptualization of those image-models in an intuition, which is in turn conceived as a complex demonstrative device in the language of thought. Furthermore, I am willing to accept this account as an adequate reading of Kant’s conception of schematization in its a posteriori guise. While I have defended these claims at length in other publications,²³ here I will take these parts of Sellars’s interpretation largely for granted – both as philosophically enlightening and as exegetically adequate.

5.3 Synthesis

With this elucidation of the schematizing activity of the productive imagination in mind let me now turn to the second task at hand: Kant’s theory of *synthesis*. In the first edition of the *CPR* from 1781, Kant introduces a synthesis of apprehension as the first part of a “threefold synthesis” (A97) that not only requires apprehension, but also a synthesis of reproduction and a synthesis of recognition. The concept of apprehension thus helps differentiate a complex synthetic process into its different aspects, which each have to be taken into account in the explication of any synthesis – empirical or a priori – that can be ‘brought to concepts’ (cf. A78/B103) and thus give us conscious representations of objects. In this complex ‘threefold synthesis’ the imagination takes up into consciousness the sensibly given material and restructures it in accordance both with the forms of intuitions specific to our (human) sensibility (space and time) and in accordance with the categories of the understanding (made suitable for our sensibility or, as Kant puts it, schematized). This threefold synthesis is thus Kant’s term in the A-Deduction for the process that gives us both image-models and intuitions in, I would argue, much the same way as outlined above. In the second edition from 1787 this product of a synthesis brought to concepts is called a *combination* (*Verbindung*):

But the concept of combination includes, besides the concept of the manifold and of its synthesis, also the concept of the unity of the manifold. Combination is representation of the synthetic unity of the manifold. (B130/1)

Combination, as a unity of synthesis according to the first edition, requires each aspect of the threefold synthesis and, in particular, the synthesis of recognition that unites the other syntheses by means of concepts. In the second edition, however, this combination is achieved by the synthesis of apprehension *alone* (guided by apperception), which is now defined as “that combination [*Zusammensetzung*] of the

²³ Cf., for instance, Haag (2007, Ch. 7 and 8).

manifold in an actual empirical intuition, whereby perception, that is, empirical consciousness of the intuition (as appearance), is possible” (B160).

The role of the concept of a synthesis of apprehension in the second edition has therefore changed in two decisive ways. Firstly, it is broader since it now covers the whole synthetic process leading to combination, rather than just the first part of a synthetic process that has to be supplemented by a synthesis of reproduction and a synthesis of recognition, as in the ‘threefold synthesis’ of the first edition. Hence, it facilitates by itself consciousness of the intuition *qua* appearance, i.e., of a representation taken as an object rather than a representation taken as a representation.²⁴ While, secondly, at the same time, its scope of possible application is significantly diminished since this synthesis is now explicitly restricted to *empirical* synthesis.²⁵ The second edition’s synthesis of apprehension takes up impressions and synthesizes them into perceptions, i.e., conscious representations.²⁶

In what follows I will be concerned mainly with the synthesis of apprehension of the first edition, i.e., the synthesis of apprehension as either an a priori or empirical first step in a complex threefold synthesis, the entirety of which alone can afford us conscious experience.²⁷ I will, accordingly, restrict my use of the term synthesis of apprehension to Kant’s use in the first edition unless otherwise indicated. Only when speaking again at the general level of observation (with which I first began) will I use the term “apprehension” in the broad sense required for characterizing the imagination as a faculty of presentation *and* apprehension, which implies the entire un-intentional activity of imagination-cum-understanding we are now about to investigate in some detail.

However, even if for the purposes of this elucidation we concentrate on the narrower concept of the first edition, we encounter quasi-definitional claims that do not fit very well with the other elements of the threefold synthesis. The problem becomes apparent as soon as we ask what exactly the role of the synthesis of apprehension is within the threefold synthesis. Kant, by way of maintaining the necessity of this aspect of the threefold synthesis, writes:

Every intuition contains in itself a manifold which can be represented as a manifold only insofar as the mind distinguishes the time in its sequence of one impression upon another; for each representation, in so far as it is contained in a single moment, can never be anything but absolute unity. In order that unity of intuition may arise out of this manifold (as is required in the representation in space) it must first be run through, and held together. This act I name the synthesis of apprehension because it is directed immediately upon intuition, which does indeed offer a manifold, but a manifold, which can never be brought about [bewirkt; J.H.] as such, and as contained in a single representation, save in virtue of such a synthesis. (A99)

²⁴ Kant, through the addition of “as appearance”, directs attention to what we have learned to call the act/object-ambiguity.

²⁵ This is made entirely clear in B162 Fn.

²⁶ Cf. A320/B377.

²⁷ What this concept is designed to cover, nevertheless, has not simply become obsolete in the second edition. It should be taken as belonging to the ‘loss’ Kant refers to in the *Preface* to the second edition concerning which he explicitly refers the reader back to the first edition. Cf. B XLII.

This quote contains many important clues and one inaccuracy. Firstly, the manifold, which every intuition contains in the way outlined in the first quoted sentence, is the manifold of an intuition that is already subject to a synthesis. Kant is asking for the conditions of the possibility of a given sensory consciousness of a manifold. He is not indicating a genetic starting point for the synthesis under investigation, in which case he would have to refer to the manifold as (synoptically) presented by receptivity.

Secondly, only if those impressions are ordered in time will we potentially distinguish the manifold as such. I say ‘potentially’ since this has to be understood as a necessary but not a sufficient condition for the distinctive consciousness of a manifold as a manifold. Other conditions have to be fulfilled, notably the successful synthesis of reproduction and recognition, i.e., the two other parts of the ‘threefold synthesis’.

Thirdly, if they were not ordered in this way (i.e., if there was no *succession* of moments but only single moments), then every one of these moments would contain only such representations as constitute an *absolute unity*. This notion of an absolute unity refers back to the reason Kant gave for the necessity of a spontaneous synthesis, namely, that otherwise no knowledge (*Erkenntnis*) would be possible, since in this case “each representation were completely foreign to every other, standing apart in isolation” (A97). It will be necessary to come back to this topic of absolute unity, however it should be clear already that such unity cannot be the unity of a *synthetic complex*.²⁸ The only remaining alternative would thus seem to be that this unity is a unity that is absolute because it is *not complex at all*, i.e., a representation that is not itself the representation of a manifold, although it must be able to become *part* of such a representation.

Fourthly, a definite complex representation is the intuition that is supposed to unite these representations *qua* absolute unities into a unitary representation of this manifold *as a manifold*. Although not every intuition is obviously a representation of a complex, every intuition itself must consequently be a complex representation. Fifthly, this unity of a complex representation is to be secured by the act of running through (*Durchlaufen*) and the holding together (*Zusammennehmung*) of the manifold. And here we run into the exegetical difficulty mentioned above: the synthesis of apprehension in this quote is assigned two different tasks – running through and holding together a given manifold – but only *one*, namely, the running through, ultimately defines its role in the threefold synthesis.

²⁸ Could it be, alternatively, a unity of a merely *synoptic complex* provided for by the purely receptive “synopsis of sense” (A97) (cf. below Sect. 5.5)? This is also implausible because it would mean a duplication of structures that is superfluous at best and incoherent at worst: we would have synoptically structured complexes *embedded* as isolated parts in synthetically structured complexes. In this case a meaningful relation between the forms of receptivity and the forms of intuition would be lost: the hypostasis of structured receptive input would amount to rational psychology, which is not in any way transcendently founded. Moreover, the argument that justifies the introduction of forms of receptivity connects them to forms of intuition and we would not have *these* forms of intuition unless we had *these* forms of receptivity (whatever they might be). Without them, this connection would get completely lost, which would be fatal to the overall transcendental justification of the picture. For more on the topic of synopsis cf. below 5.5.

5.4 A ‘Threefold Synthesis’

This claim about the restricted role of the synthesis of apprehension as part of the threefold synthesis can be substantiated by investigating the further deployment of the synthesis of apprehension in the A-Deduction and its relation to the other parts of this complex synthetic process. As Kant makes clear in his “deduction from below” (in A 119 ff.), this synthesis is a function of imagination, an “active faculty for the synthesis” (A 120) that is “immediately directed upon perceptions” (ibid.). And he continues saying: “Since imagination has to bring the manifold of intuition into the form of an image, it must previously have taken the impressions up into its activity, that is, have apprehended them.” (ibid.)

This taking up of impressions into the activity of spontaneity is the task of the synthesis of apprehension and is explicitly distinguished from the generation of a “connection of the impressions [Zusammenhang der Eindrücke]” (A121): the first operation corresponding to the running through, the second to the holding together of the first definition of the synthesis of apprehension. For this second operation we consequently need a further activity, this time the “*reproductive* faculty of imagination” (ibid.; emphasis added), that is, the synthesis of reproduction. But before something can be reproduced, Kant seems to reason, it must be taken up into the faculty that does the reproducing. Those two synthetic steps are indeed “inseparably bound up with” (A102) each other so intimately that together they constitute the “transcendental faculty of imagination” (ibid.). However, they can be abstractly distinguished as different aspects of one comprehensive synthesis.

Yet even then the synthesis remains incomplete: a further step, making the synthesis truly ‘threefold’, is needed. This would be the *synthesis of recognition*. It answers a problem on which Kant elaborates in the ‘Deduction from below’:

If, however, representations reproduced one another in any order, just as they happened to come together, this would not lead to any determinate connection of them, but only to accidental collocations [Haufen]; and so would not give rise to any knowledge. Their reproduction must, therefore, conform to a *rule*, in accordance with which a representation connects in the imagination with some one representation in preference to another. (A121; emphasis JH)

To provide us with objective representations, i.e., knowledge (*cognitio*), the synthesis must reproductively synthesize according to a rule. This rule determines how the apprehended and reproduced representations should be ‘brought into an image’ (cf. A120). According to my Sellarsian picture, the rule determines how the apprehended representations are united into image-models that serve as a reference for the corresponding intuitive representation of an object of experience. This rule is therefore nothing else than the empirical concept (containing the schema) that guides the synthetic activity of the productive imagination. And the aspect of the threefold synthesis that ensures its being executed according to a rule is the synthesis of recognition.²⁹

²⁹ When introducing the synthesis of recognition Kant puts this in terms of the synthesis guaranteeing the identity of a synthesized object. He phrases the solution to the problem from A 120 explicitly in terms of the synthesis of recognition in A124. For more on the synthesis of recognition cf. Haag (2007, 220 ff).

This aspect makes clear why Kant can refer to the understanding in one (prominent) place as a “faculty of rules” (A126). However, this part of the threefold synthesis may seem less inseparably bound up with the other two in the sense that it might be possible for there to be non-rational creatures with a faculty of imagination that comprises only the first two aspects (depending on their additional ability for associative connection), whereas creatures capable of apprehending only in the sense under discussion would not be able to find their way around in the world.³⁰ At any rate, for finite rational beings such as us, the synthesis *has* to be threefold. Otherwise we would not have intuitions of objects of experience, i.e., objects that are synthesized in accordance with schematized categories, as Kant aims to prove in the ‘deduction from below’ – the culminating part of the A-Deduction.

These observations concerning the inseparability of the three aspects of the threefold synthesis point to another feature of its first part: the synthesis of apprehension alone cannot, through simply taking up our impressions into the activity of the imagination, generate *conscious* representations (*perceptiones*). For a representation to be conscious more is needed, as Kant forcefully contends in the argument from associability in A121/2 where he logically ties consciousness to the possibility of apperceptive consciousness and the related synthesis of recognition.³¹ Here,

Rosenberg claims that the synthesis of apprehension alone could produce image-models, while the synthesis of recognition is only needed to afford us intuitive representations in Sellars’s sense (Rosenberg 2007b, 240). The third aspect of the threefold synthesis, the synthesis of recognition, in his view, is reserved for “perception across time” (ibid.). (Kant indeed does talk about identity through time in this context. But identifying this with the task of recognition amounts to confusing a particular argument for the synthesis of recognition with the description of its contribution to the threefold synthesis.)

Furthermore, Rosenberg’s account seems misguided in a number of other ways: the synthesis of apprehension is nothing more than the taking up into consciousness of the manifold of receptivity; for the construction of image-models the impressions must be reproduced. Otherwise we would have only isolated sense-impressions. (Cf. Kant’s example of the drawing of a line in thought in A102.) The synthesis of recognition likewise cannot be omitted in the generation of image-models. It is needed for executing the recipes for construction, that are not part of the intuition of an object of experience, say, of a dog, though they *are* part of the corresponding concept of *dog*. (This last remark refers back to my repudiation of Rosenberg’s claim that the recipe (schema) used in the generation of the image-model should simply be *identified* with the “demonstrative conceptualization” (Rosenberg 2007a, 273), that is, the intuition.)

³⁰ Cf. the letter to Marcus Herz from May 26, 1789. McDowell quotes this letter in support of his claim that “our sensibility should be something non-rational animals also have” (McDowell 2009, 117). Accordingly, this claim is introduced by McDowell as a constraint for any successful account of sensory consciousness.

³¹ In brief, I take Kant to argue that representations could not be conscious unless they are associable. For representations can be only conscious through their combination into a complex. Associability, however, is the presupposition of such a combination. Conscious representation, hence, presupposes associability and cannot be thought without it – which must, consequently, be conceived as a condition of the possibility of conscious representation. To this end, however, it has to be provided with an objective, not merely subjective ground. If associability is a purely subjective ground it has to be supplemented by a corresponding ‘affinity of appearances’ (A122) that turns out to be guaranteed by the unity of apperception, i.e., self-consciousness. Cf. Haag (2007, 241–247). The relationship between apperception and (empirical) synthesis of recognition is discussed in *ibid.*, 239 f.

apprehended representations become conscious only *as constituent parts of an intuition*, which implies that the threefold synthesis must complete the synthesis of complex representations (intuitions) before *any* consciousness of the synthesized parts becomes possible. Only then can we, in a further *abstractive* step,³² become conscious of the parts that constitute this unitary manifold *as parts* of a manifold.

If this consideration is correct, it will follow that the threefold synthesis itself is merely a necessary, but not a sufficient condition of the *consciousness of the sensible aspects* of an intuition as such, even if it would of course be a necessary and sufficient condition for *sensory consciousness*. Furthermore, it would put into question the possibility of there being non-rational creatures, at least in the way brought up above (i.e., creatures limited to the first two steps of the threefold synthesis).³³ Consequently, the ascription of consciousness to non-rational *animals* Kant allows for in the letter to Herz from May 1789 should be read as an ascription on a level of abstraction *different* from the one at which his discussion of the threefold synthesis is located. It is an analogous measure that presupposes conceptually synthesized representations, and therefore cannot elucidate the conceptual means used in the description of their synthetic constitution.³⁴

5.5 The Synopsis of Sense

If the task of the synthesis of apprehension was the taking up of impressions in the activity of imagination, why is it still called a synthesis, which originally was introduced as an “act of putting different representations together” (A 77 / B 103)? This may not appear so strange if we remember that the synthesis of apprehension is merely one *aspect* of a complex synthetic act. The denomination thus would reflect its being part of a synthetic act and highlight the fact that it is the imagination (the faculty of synthesis) that does the apprehending. We can perhaps shed further light on this denomination if we turn to two closely related questions of crucial importance for the discussion of sensory consciousness and its conceptual shaping: (1) What are the impressions taken up into the activity of the imagination and (2) are they changed in undergoing this procedure?

³² I think that this abstraction would have to be very much like what McDowell suggests in 2009, 119–122. What this analysis deliberately leaves out, of course, is the alternative concept of a *receptive* sensation. If the present account of synthesis of apprehension should turn out to be correct, receptive sensations obviously would not be accessible to a comparable abstractive approach.

³³ This should serve as a powerful reminder of the fact that the metaphorical talk about ‘steps’, while sometimes necessary for elucidating the details of the threefold synthesis, is ultimately misleading, since it suggests that each step could take place independently of the following. To call the three parts of this synthesis ‘aspects’ is more adequate.

³⁴ This is meant as a clarification of my assessment of this subject matter in Haag (2007, 231 Fn). If correct, it will count against the use McDowell makes of the continuity between the sensory consciousness of animals and rational beings.

The first part of the question has to take into account the synopsis of sense that I already mentioned in passing. What is this synopsis? It is a concept that appears only in the A-Deduction, where it is sharply distinguished from any kind of synthesis.³⁵ Synthesis is always a function of the imagination (i.e., a function of the understanding taken in a certain way), whereas synopsis does not require any activity of the imagination and hence no synthetic activity. Synopsis is a function of sense and sense without synthesis is sense without spontaneity, in other words, sense as sheer receptivity. Synopsis is, therefore, a function of sheer receptivity.

An exact exegetical analysis of the background of the frequently neglected³⁶ concept of synopsis is beyond the scope of this paper,³⁷ however, I must at least sketch some considerations relevant in this context. This is important because the synopsis of sense proves to be the key ingredient in a conception of conceptually shaped empirical consciousness that guarantees that the conceptual normativity of this process is duly guided by nature ‘from without’. While synopsis is not a synthesis, it nonetheless involves some sort of structuring of given sensory material, as Kant makes clear in the only other mention of synopsis shortly afterwards:

If each representation were completely foreign to every other, standing apart in isolation, no such thing as knowledge would ever arise. For knowledge is a whole in which representations stand compared and connected. As sense contains a manifold in its intuition, I ascribe to it a synopsis. But to such synopsis a synthesis must always correspond; receptivity can make knowledge possible only when combined with spontaneity. (A97)

If synopsis *corresponded* to synthesis with respect to its being a faculty that unites otherwise distinct and isolated representations, we could certainly conclude that some order is already imposed on the manifold ‘in intuition’ by the corresponding synopsis of sense.

Yet, as we learn from the *Amphiboly of the Concepts of Reflection* of the CPR, the structuring in this case – while (like all structuring) it has to be an instance of determination and determinable or, as Kant puts it, form and matter – it can also take place in an order in which form, unlike *conceptual* form, *precedes* matter. In what sense can determination precede the determinable, i.e., form precede matter? Kant elucidates this with admirable clarity in the *Eberhard-Streitschrift*, writing that the characteristic receptivity (*eigentümliche Rezeptivität*) of the mind (*Gemüth*) is an innate disposition to receive sensory affection.³⁸ It is the mere disposition to react (in no sense spontaneously but merely passively) when acted upon and affected by things-in-themselves. This is the sense in which even the synopsis of sense can be a priori, as Kant puts it in A94. Synopsis of sense would thus be understood, alternatively, as the process of structuring or its result.³⁹

³⁵ Cf. A94.

³⁶ An exception is Waxman (1991, 228–225). Waxman interestingly sees as part of the role of synopsis the purgation from form that I will ascribe to the Synthesis of Apprehension.

³⁷ But compare my attempt in Haag (2007, Ch. 4).

³⁸ Cf. 8:222.

³⁹ The concept of synopsis shares this particular ambiguity with the concept of synthesis that can be used (and is used by Kant) likewise to alternatively represent the process or the product of synthesis.

Although, as we have seen, in the faculty of sense form precedes matter, I now want to turn first to the material aspect of this activity in sheer receptivity. The first question raised by Kant's characterization of synopsis certainly concerns the nature of the manifold that sense receptively orders through its synopsis, i.e., its *affective input*. Hardly anything more can be said about this manifold than that our receptivity is affected by it. To say more would compromise Kant's strict opposition to the possibility of knowledge of the nature of the things-in-themselves, by which we are affected in non-empirical affection. All we can accordingly know is that we have to think this original manifold of affection for transcendental-philosophical reasons related to guidedness and passivity. Thus, in cases like this, it is justified to claim *that* such an affective input has to be assumed since it is a condition of the possibility of experience: without it, we would lose a necessary sense of guidance from without. Nonetheless, we are not entitled to make *any* meaningful assumptions as to the *nature* of this affective input.

Very similar things must be said about the *products of the synopsis of sense*, which Kant calls "impressions" (A120; A121) in the 'Deduction from below', referring to the manifold of sense *preceding* any synthesis. We can know nothing about the nature of these impressions, as this would mean knowing something substantial (as opposed to purely formal or transcendental) about something that is essentially non-synthesized. We know, however, that we do in fact have such synoptically structured impressions presenting a manifold "for intuition" (B145) that has to be "given *prior to* any synthesis of understanding and *independent of it*" (ibid.; my emphasis). As these quotes indicate, the purely receptive input is not completely left out of the second edition. Earlier in its *Transcendental Deduction* Kant even takes up – what in the first edition of the *Critique of Pure Reason* would have been – the forms of receptivity and the products of the synopsis of sense. At the very beginning of §15, he writes:

The manifold of representation can be given in an intuition which is purely sensible, that is, nothing but receptivity; and the form of this intuition can lie a priori in our faculty of representation, without being anything more than the mode in which the subject is affected. (B129)

Kant's picture of sensory consciousness therefore implies the existence of completely non-synthetic, non-spontaneous and a fortiori neither conceptual nor intentional sensory structured *material*, namely, the *synoptically structured impressions of sheer receptivity*. With Kant we can distinguish a *manifold for intuition*⁴⁰ from an *intuition of a manifold*, i.e., a manifold *not for intuition* (as the material on which the synthesis of productive imagination can operate) but a conscious manifold *for us qua subjects of experience*. Notice that the products of the synopsis would be sensations completely located below the line that separates the realm of spontaneity from that of sheer receptivity. As such, they cannot be structured by space and time as forms of intuition in the sense elucidated above, namely, as themselves the products of an a priori synthesis. To suggest they were would be tantamount to saying that the

⁴⁰ B145.

result of this supposed structuring through receptivity was a structuring through spontaneity after all. It seems, accordingly, that the affective input would be structured by what we might along with Sellars call *forms of receptivity* as opposed to the *forms of intuition* that form the synthetic process.⁴¹

The first part of our above question concerning the nature of the impressions taken up into the activity of the imagination via the synthesis of apprehension therefore has a quite straightforward answer: the items taken up can only be the synoptically structured, purely receptive sensations. If this is correct, it will imply an answer to the second part of the question concerning whether those impressions are changed in undergoing this procedure? For, what is taken up into the activity of the imagination by the synthesis of apprehension *cannot* survive this process or, to put it more cautiously, at least cannot survive it unchanged. We settled earlier⁴² on a picture of the *absolute unity* of the products of this taking up as being not at all complex, i.e., as being themselves representations that are not the representation of a manifold. It is precisely in this spirit that Kant writes:

Apprehension by means merely of sensation occupies only an instant, if, that is, I do not take into account the succession of different sensations. As sensation is that element in appearance the apprehension of which does not involve a successive synthesis ..., it has no extensive magnitude. (A167/B209)⁴³

The receptive sensations, in being taken up into the activity of the productive imagination, therefore lose the structure they had eventually gained through the synopsis of sense, i.e., by means of the forms of receptivity. It is an important aspect of the synthesis of apprehension that it purges the receptive sensations of this structure. The resulting sensations are thus already a product of spontaneity. They are, in this sense, *spontaneous sensations*. It follows that we have a second kind of sensation that must be distinguished carefully from the products of mere receptivity: the concept of sensation as the spontaneous product of the synthesis of apprehension.

But why do these sensations have to be of *absolute unity* in the first place? Some authors (including, in all likelihood, Sellars)⁴⁴ have attributed this claim to Kant's adherence to the sensualistic proclivities of his time. But this interpretation misses the point, neglecting the interplay of form and matter (a Kantian principle of reflection⁴⁵) at this stage of the transcendental reflection: sensations can be neither temporally nor spatially complex since they must serve as matter for the forms that structure them. To be complex at this level of analysis would *mean* to have temporal or spatial extension. This then is the reason for their occupying 'only an instant': they do not have a temporal or spatial extension because otherwise they would be

⁴¹ Cf. Sellars (1968, 29). Sellars, of course, criticizes Kant for not paying attention to this very distinction. If my interpretation is correct, however, in this context Sellars fails to do full justice to the subtlety of Kant's approach.

⁴² Cf. p. 10 above.

⁴³ Cf. A145/B184, 16:662, 18:268.

⁴⁴ Cf. Sellars (1968, 27) and Henrich (1976, 17).

⁴⁵ Cf. A266f./B322f.

already formed. Rather, they are simple or atomistic states that contain exactly one monadic sensory quality.

It is important to notice that the same cannot be said for the *receptive sensations* as they are taken up in spontaneity. For even after they have lost their purely receptive synoptic structure, they themselves might have properties stemming from their being affections of our receptivity by things-in-themselves. Due to the inaccessibility of unsynthesized reality, i.e., reality-in-itself, we can neither rule out nor confirm that they are already structured complexes, which have only gained a further layer of structure from the forms of receptivity. The *spontaneous* sensations, on the other hand, should be identified with the concept of sensation that Kant refers to in the ‘Stufenleiter’ as a conscious representation (perception) “which relates solely to the subject as the modification of its state” (A320/B376).

This characterization is in line with Kant’s first introduction of the concept of apprehension at the very beginning of the A-Deduction, which classifies products of apprehension as “modifications of the mind in intuition” (A97). And this again seems to take up the theme from the *Transcendental Aesthetics* where colors, as paradigmatic sensations, are classified as being “not properties of bodies to the intuition of which they are attached, but only modifications of the sense of sight, which is affected in a certain manner by light” (A28) and sharply distinguished from space, which “as condition of outer objects, necessarily belongs to their appearance or intuition” (ibid.).

And yet there are important conceptual differences: we have here, I would like to suggest, two different *ways of conceiving* what – ontologically speaking – is ultimately one and the same entity. The underlying ontology is the ontology of sensations as modifications of *empirical* subjects (whereas receptive sensations are to be thought of as modifications of transcendental subjects). In the *Transcendental Deduction*, these modifications of the empirical subject are considered from the perspective of transcendental-philosophy, while in the quote from the *Aesthetics* they are viewed from the perspective of natural science mechanistically conceived.

In *transcendental perspective* those modifications serve the purpose of providing matter, which, via forms of intuition, schematized categories and empirical concepts, can be formed into image-models that can then be taken (in an intuition) as objects of experience. From the *perspective of natural science* they are the result of the affection of the senses of an empirical subject brought about by the purely spatio-temporally extended objects of empirical reality – an affection that can be completely described by means of mechanistic natural science (the successor concept of matter *sub specie scientia naturalis* would then be “force as the concept of mechanics” (Sellars 1968, 45).⁴⁶)

Given the discussion of the concept of an image-model and its difference from the concept of intuition (namely, image-models being perspectival arrangements of sensibly given material according to a recipe, while intuitions are the taking of those

⁴⁶ Sellars simply identifies this scientific concept of color as secondary quality with the transcendental concept of “the real which is an object of sensation” (B 207), which, if there is something to my considerations, amounts to a confusion.

very models as the objects to which we demonstratively refer), there are, strictly speaking, *two* transcendental conceptions of empirical sensation that are blended in this characterization: one as matter in appearances (i.e., as matter determined through forms of intuition) and another as matter for a represented object. The first would come into play in the construction of image-models, the second in the intuitive reference to objects of experience.

Be that as it may, the fact remains that *any* class of sensations (given it has a certain manifold) could fulfill the role of matter in this context: unlike the forms of intuition (space and time), the concrete character of sensation is completely arbitrary. From the transcendental perspective we only need *some* matter – *whatever* quality it might encompass.⁴⁷ Those modifications can therefore play their transcendental role while being nothing but largely arbitrary modifications of empirical subjects. We can conceive of them *either* (taking a transcendental viewpoint) as a result of the synthetic taking-up of the structured sensations of receptivity by the transcendental subject *or* (in describing empirical reality *as* reality) as produced by the empirical subject's being affected by objects that are located in space and time (just like ourselves *qua* empirical subjects).

In both perspectives, however, the sensations would, ontologically speaking, be modifications of the *empirical* subject. For even from the transcendental viewpoint they are conceived as modifications of the subject taken as an empirical subject (since they would be conscious representations and there can be no consciousness of the modifications of the transcendental subject). As modifications of a subject, conscious representations must be appearances, whereas the existence of the conscious subject is not mere appearance.⁴⁸

Where does that leave us concerning the *continuity* between the sensations of receptivity and the sensations of spontaneity? Talk of the synthesized sensations being the *same sensations* as the receptive sensations turns out to be profoundly inadequate for a number of reasons. Firstly, we cannot know anything about the existence of a meaningful continuity between the two kinds of sensation due to the inaccessibility of the transcendental subject's modifications as they are in themselves. Secondly, there need not be any continuity concerning their qualitative content between receptive sensations and the products of the synthesis of apprehension, which are the result of the completely arbitrary subjective disposition of an empirical subject. And, thirdly, since sensations that are products of this kind of synthesis are not in any way formally structured, continuity in form is likewise excluded.

Continuity in form being excluded, qualitative continuity remains completely arbitrary and ultimately unnecessary. Moreover, given the items that could answer this question are inaccessible, there seems to be no meaningful way left of talking about a continuity between both kinds of sensations: we cannot say that spontaneous sensations are sensations of receptivity taken as something else. Whatever the

⁴⁷ Sellars, of course, would deny just that, since he sees space and color as inextricably bound up with each other. But Kant has an interesting argument against this to the effect that we only need *some* matter, but exactly *this* kind of form. For a discussion compare Haag (2007, 142–150).

⁴⁸ Cf. B157/8.

material basis for the synthesis of apprehension, it is changed so radically that there is no way back from synthesized sensation to receptive sensation – not even by (Sellarsian) analogy.

5.6 Synthesis a Priori and the Concept of Guidance

But is this separation not too strict? Through the synthetic process that leads to the construction of image-models and, in the last instance, objects of experience, we certainly have provided for the normative-conceptual structuring of the world-as-perceived. But, in so doing, are we not losing all philosophically interesting contact with the happenings below the line that separates spontaneity and receptivity in this way? Will not, in other words, the loss of a meaningful way of talking about the *same* sensations be linked with the loss of any meaningful concept of guidance from without?

To find a way to prevent this, we must invoke the a priori dimension of the synthesis of apprehension and, with it, the formal aspects above *and* below this line. In the first edition's *Transcendental Deduction* Kant is very explicit about the a priori function of the synthesis of apprehension. Without such a synthesis, he writes, "we should never have a priori the representations either of space or time" (A99). And he continues: "They can be produced only through the synthesis of the manifold which sensibility presents in its original receptivity. We have thus a pure synthesis of apprehension." (A99/100)

Without such a synthesis we would have no relation to what is presented through 'original receptivity' to the subject of experience and hence no consciousness of either space or time. In the words of the B-Deduction: we would have "forms of intuition" (B160 fn.), but no "formal intuition" (ibid.) of space and time. I interpret Kant's distinction between the forms of intuition and formal intuition as corresponding to my Sellarsian distinction between forms of receptivity and forms of intuition respectively.⁴⁹

The synthesis of apprehension, accordingly, must have an a priori use, since our forms of intuition are a priori forms. In this pure synthesis a manifold is somehow synthesized into representations of space and time.⁵⁰ Let me now give a brief sketch of this process of a priori synthesis since the way it ought to be conceived will serve to provide a way out of the problem outlined above of our losing contact with the 'below the line' elements of sensory consciousness.

First of all, it must be acknowledged that the synthetic process cannot provide us with intuitive representations of space and time unless we expand it (in the spirit of

⁴⁹ Since Kant often uses form of intuition for what he calls formal intuition in B160 fn., I have decided not to use his somewhat confusing terminology in this case, but rather to stick to Sellars's terminological suggestion instead.

⁵⁰ These (undetermined) *intuitive* representations of space and time have to be distinguished carefully from the (determined) *concepts* of space and time based upon them. Cf. e.g. A25/B39.

the B-Deduction) to encompass the entirety of the threefold synthesis – or, at the very least, apprehension and reproduction must be taken together.⁵¹ Furthermore, we would also seem to need the synthesis of recognition, as the synthesized representations would then be conscious representations. This is, however, not a major problem since we have to think the aspects of the threefold synthesis together anyway.

The problem with an a priori synthesis of a manifold is that it is not easy to see on what kind of *matter* this synthesis should operate. We do not have any formal manifold in the sense of a given manifold of forms.⁵² The forms of receptivity are not given, save by their product, the receptive sensations. The only manifold that provides material for a synthesis of apprehension is therefore the formed material that is given to our spontaneity by the synopsis of sense – a thoroughly *a posteriori* product. How can we get from there to a pure synthesis of an a priori manifold? The solution seems to lie in a use of *abstraction* that, as it were, purifies the synoptically given manifold by concentrating on the process of synthesis and neglecting its material content. Kant illustrates this concentration on the manner or mode in which the material is given in apprehension in § 24 of the B-Deduction:

We cannot represent the three dimensions of space save by setting three lines at right angles to one another from the same point. Even time itself we cannot represent, save in so far as we attend, in the drawing of a straight line (which has to serve as the outer figurative representation of time), merely to the act of the synthesis of the manifold whereby we successively determine inner sense, and in so doing attend to the succession of this determination in inner sense. Motion, as an act of a subject (not as a determination of an object), and therefore the synthesis of the manifold in space, first produces the concept of succession – if we abstract from this manifold and attend solely to the act through which we determine the inner sense according to its form. (B154)

What Kant describes here is not an abstract separation of the formal and qualitative content of a given representation, but rather a concentration on the procedural character of the act of synthesis. That he talks about the concept of succession and not about time as a form of intuition (formal intuition) should not deter us from the application of the method described to the intuition of time (or space), since this undetermined intuitive representation, even on Kant's description, would be a necessary precondition of the determined concept of time as succession.

In the act of synthesis we therefore have to concentrate on the properties that belong solely to the formal aspects of sensibility. To this end, we abstract from the qualitative aspects apprehended in the synthesis as well as from those formal aspects that do not belong to sensibility but solely to understanding (and which are, consequently, not aspects of the act that gives us forms of intuition but rather a priori concepts). If we proceed in this way, we are left with what amounts to formal intuitions of time and space. This abstractive account is therefore able to explain

⁵¹ Which could, by the way, help to explain the lapse on Kant's part in the original characterization of the synthesis of apprehension: he might have been preoccupied with a priori synthesis.

⁵² Vaihinger tried to defend this line of interpretation. R.P. Wolff interpreted Kant along similar lines, but correctly judged it a failure – unfortunately, however, as a failure on Kant's part, and not on the part of his interpretation. Cf. Vaihinger (1892, 102–111) and Wolff (1963, 218–223).

why we come to have pure intuitive representations of space and time only *after* using them as forms of intuition in the schematizing of the categories in the modeling of image-models. These intuitions are a priori since we understand through the act of abstraction that they are the forms that underlie each and every intuitive representation. Otherwise we could not have obtained a consciousness of them through this kind of process.

How is this a priori synthesis of apprehension supposed to restore the contact to the synoptical elements below the line that separates receptivity from spontaneity – a contact that guarantees the ‘guidance from without’ required by the conceptual order? Since the qualitative aspects of our sensory consciousness above and below the line cannot be used for this purpose, we should now heed its formal aspects not only above the line but also below it. I have noted before that it is not possible to give any direct characterization of these forms. But now we can substantiate the previously empty, indirect characterization of those forms of receptivity as *being of a kind that produces synoptically structured representations, which are able to serve as the basis for a synthesis producing exactly our forms of intuition*. In other words, we would not have synthesized exactly *these* forms of intuition by a synthesis that is both a priori *and* transcendental unless we had been given material synoptically *formed like this* “manifold which sensibility presents in its *original* receptivity” (A100; my emphasis).

Though the inaccessibility of the below the line aspects is guaranteed in this description of the relation of formal aspects above and below that line which separates sheer receptivity from spontaneity, the contact between the two pairs of forms (forms of receptivity and forms of intuition) is of a kind completely different from the relation between the receptive and spontaneous sensations. This is mainly due to the fact that the transcendental status of forms of intuition sharply distinguishes them from objective sensations. There is a transcendental argument for the *existence* of sensations (as matter of experiences), but there is no such argument as to their concrete nature. Yet, in the case of forms of intuition, there *is*, in Kant’s opinion, such an argument (e.g. in the case of space, from the necessity of thinking something distinct from us).⁵³ It is therefore exactly *those* forms of intuition that we have to synthesize in order to be able to intuitively refer to objects of experience. And this fact is explained by our sensibility being characterized by exactly *these* forms of receptivity.

Those representations then, which we *de facto* entertain in perception of empirical reality, are representations dependent on the synoptically given material that is structured by the forms of receptivity in their individual formal aspects – their individual shape and their individual duration. Furthermore, what is structured in this process is in turn due to our being affected by things-in-themselves. Given both these claims, it seems reasonable to conclude that the idea of a conceptually, and hence normatively, laden process being ‘guided from without’ is done justice in the Kantian system.

⁵³ Cf. Haag (2007, 142–150).

References

- Haag, Johannes. 2007. *Erfahrung und Gegenstand. Das Verhältnis von Sinnlichkeit und Verstand*. Frankfurt a.M: Vittorio Klostermann.
- Henrich, Dieter. 1976. *Identität und Objektivität. Eine Untersuchung über Kants transzendente Deduktion*. Heidelberg: Winter.
- Kant's works are quoted in the usual manner after *Kants gesammelte Schriften*, vols. I–XXII ed. Preußischen Akademie der Wissenschaften, Berlin 1902 ff., vol. XXIII ed. Deutschen Akademie der Wissenschaften, Berlin, 1956, vols. XXIV–XXIX ed. Akademie der Wissenschaften zu Göttingen, Berlin 1966 ff. The translation mostly follows the *Cambridge edition of the works of Immanuel Kant*, eds. P. Guyer and Allen W. Wood.
- McDowell, John. 2009. Sensory consciousness in Kant and Sellars. In *Having the world in view. Essays on Kant, Hegel, and Sellars*, ed. John McDowell, 108–126. Cambridge, MA: Harvard University Press.
- Rosenberg, Jay. 2007a. Divergent intuitions: McDowell's Kant and Sellars's Kant. In *Wilfrid Sellars: Fusing the images*, ed. Jay Rosenberg, 266–290. Oxford: Oxford University Press.
- Rosenberg, Jay. 2007b. Perception vs. inner sense: A problem about direct awareness. In *Wilfrid Sellars: Fusing the images*, ed. Jay Rosenberg, 230–242. Oxford: Oxford University Press.
- Sellars, Wilfrid. 1963. Phenomenalism. In *Science, perception and reality*, ed. Wilfrid Sellars, 60–105. London: Routledge and Kegan Paul.
- Sellars, Wilfrid. 1968. *Science and metaphysics. Variations on Kantian themes*. London: Routledge & K. Paul.
- Sellars, Wilfrid. 1976. Kant's transcendental idealism. *Collections of Philosophy* 6: 165–181.
- Sellars, Wilfrid. 1978a. The role of imagination in Kant's theory of experience. In *Categories: A colloquium*, ed. H. Johnstone, 231–245. University Park: Pennsylvania State University.
- Sellars, Wilfrid. 1978b. Berkeley and Descartes: Reflections on the theory of ideas. In *Studies in perception. Interrelations in the history of philosophy and science*, ed. P. Machamer and R. Turnbull, 260–311. Columbus: Ohio State University Press.
- Sellars, Wilfrid. 1982. Sensa or sensings: Reflections on the ontology of perception. *Philosophical Studies* 41: 83–111.
- Vaihinger, Hans. 1892. *Commentar zu Kants Kritik der reinen Vernunft*. Stuttgart (Bd.1: 1881; Bd. 2:1892).
- Waxman, Wayne. 1991. *Kant's model of the mind*. Oxford: Oxford University Press.
- Wolff, Robert P. 1963. *Kant's theory of mental activity: A commentary on the transcendental analytic of the 'critique of pure reason'*. Cambridge, MA: Harvard University Press.

Chapter 6

Naturalized Epistemology and the Genealogy of Knowledge

Martin Kusch

6.1 Introduction

Naturalized epistemology aims to study knowledge as a natural phenomenon, or more precisely, as a *natural kind*. Some influential advocates of this view insist that such investigation must rely first and foremost on the methods and results of the natural sciences. In so doing these authors reject other forms of epistemology that focus on the *norms* governing the use of the *concept* of knowledge. This paper is an attempt to clarify the difference between these two positions and to narrow the gap between them.

My central paradigm of a theory that focuses on the concept of knowledge is Edward Craig's research programme as set out in his two books, *Knowledge and the State of Nature* (1990) and *Was wir wissen können* (1993). Several authors have tried to use and develop this programme over the past decade (Beebe 2012; Fricker 2007, 2010; Gelfert 2011; Kappel 2010; Kelp 2011; Kusch 2009, 2011; Neta 2006; Pritchard 2009; Williams 2002). Craig himself calls his approach "conceptual synthesis" or "practical explication" (1990). I prefer the label "genealogy of knowledge" that was first suggested by Bernard Williams (2002) since it better captures the *historical nature* of the approach.

In this essay I shall attempt to defend and reinterpret Craig's project in response to criticism put forward by Hilary Kornblith in his recent paper "Why Should We Care about the Concept of Knowledge?" (2011). Kornblith is the most prolific and insightful defender of "naturalized epistemology" in contemporary philosophy (1993, 2002). His criticism of Craig thus provides a welcome opportunity to explore the relationship between naturalized epistemology and the genealogy of knowledge.

M. Kusch (✉)

Department of Philosophy, Faculty of Philosophy and Education,
University of Vienna, Vienna, Austria
e-mail: martin.kusch@univie.ac.at

I want to make plausible four theses. First, Kornblith's interpretation of Craig is mistaken: Craig is not committed to denying that knowledge is a natural kind. Second, Kornblith overlooks that Craig's project has affinities with naturalized epistemology. Third, Kornblith's use of Richard Boyd's theory of natural kinds is puzzling: Boyd advocates "promiscuous realism", Kornblith rejects it. And fourth, Craig's genealogy helps to understand unity and disunity in both concepts and natural kinds of knowledge.

6.2 Kornblith's Criticism of Craig

I begin with a summary of Kornblith's criticism of Craig's genealogy. Kornblith interprets Craig as maintaining that knowledge is not a *natural kind* like water or aluminium, but an *artificial kind* like table or monarchy (2011: 43–44). Kornblith bases this reading on one central but short passage in Craig (1990):

Couldn't it be that knowledge, like water, is important stuff, and that the purpose of the concept is simply to enable us to think and talk about it? ... I am fairly confident that this is mistaken. Knowledge is not a given phenomenon, but something that we delineate by operating with a concept which we create in answer to certain needs, or in pursuit of certain ideals. The concept of water, on the other hand, is determined by the nature of water itself and our experience of it. (1990: 3)

Kornblith's own views on knowledge—in so far as they are relevant here—directly contradict the position he ascribes to Craig. As Kornblith sees it, knowledge is a natural kind since the "category" of knowledge plays a significant explanatory and predictive role in one particular natural science: *cognitive ethology* (the science of animal behaviour). In fact, Kornblith believes that we should let cognitive ethology tell us what knowledge really is. He also insists that "... the kind of knowledge that philosophers have talked about all along is just the kind of knowledge that cognitive ethologists are ... studying" (2002: 30). Moreover, just like chemists rightly ignore folk concepts of water or aluminium, so also epistemologists should pay little attention to folk concepts of knowledge or justification: "... our concepts of knowledge and justification are of no epistemological interest" (2006: 12). Our folk concepts and intuitions concerning knowledge and epistemic justification do not tell us what knowledge and justification really are. Here too the investigation must focus on the 'stuff' itself. Kornblith's conception of natural kinds is that of his teacher Richard Boyd for whom natural kinds are homeostatic clusters of properties (Kornblith 2002: 61, cf. 2007a, b, 2011; cf. Boyd 1980, 1983, 1988, 1991, 1999). Applied to the concepts of knowledge this conception suggests the following formulation:

The knowledge that members of a species embody is the locus of a homeostatic cluster of properties: true beliefs that are reliably produced, that are instrumental in the production of behavior successful in meeting biological needs and thereby implicated in the Darwinian explanation of the selective retention of traits. (2002: 62)

Kornblith rebuts six arguments that *prima facie* speak *against* his position and *in favour* of the view he attributes to Craig. Some of these arguments Kornblith finds in *Knowledge and the State of Nature*, others Kornblith invents on Craig's behalf.

According to the first argument, knowledge cannot be a natural kind since it is generated by humans. Kornblith replies that we think of water as a natural kind even though we sometimes produce it in the laboratory (2011: 44). The second argument says that knowledge cannot be a natural kind since it has an important social role. Kornblith counters by remarking that natural kinds can have social roles, too. Gold is a case in point. Moreover, the social role of gold can, at least in good part, be explained by its natural properties. One relevant such natural property is the relative rarity of gold (2011: 45). The third argument is that knowledge answers to specifically human needs. Kornblith denies this: animals have knowledge, too. The fourth argument is a variant of the third: flagging good informants is key to knowledge. Again, Kornblith responds by rejecting the idea that knowledge has a special link to humans. Animals can be said to know things, but animals do not flag good informants (2011: 46). The fifth argument amounts to the claim that the concept of knowledge is a social construct. Kornblith grants as much but opposes the inference from this claim to the idea that knowledge itself is social (2011: 47). Finally, sixth, Kornblith imagines the Craigian to emphasise that knowledge is normative—and hence not a natural kind. Kornblith accepts the premise but rejects the conclusion: health too is normative but it still is a natural kind (2011: 48).

6.3 Is Knowledge a Natural Kind?

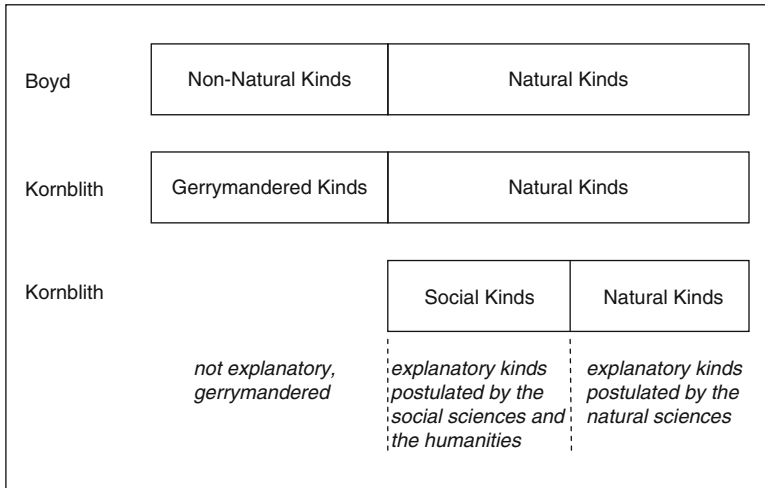
In this section I shall address the question whether knowledge is a natural kind without yet bringing in Craig’s genealogy of knowledge.

There are of course philosophers who deny that there are *any* natural kinds—never mind whether we are dealing with knowledge, water or aluminium. For instance, Ian Hacking writes that “some classifications are more natural than others, but there is no such thing as a natural kind” (2007: 203). I sympathise with this scepticism, but I shall not make use of it here. For the purposes of this paper I shall grant Kornblith the idea that there are natural kinds of some sort.

I have already mentioned that Kornblith accepts Boyd’s conception of natural kinds. Boyd defines natural kinds as causally important “homeostatic property clusters”. These are co-occurring properties such that

1. the presence of some of them tends to favour the presence of others of them; or
2. there are underlying mechanisms that tend to maintain the presence of all or most of the properties; or
3. both (i) and (ii). (Boyd 1988: 197; 1999: 143)

Boyd denies that there is *one unique set* of natural kinds. That is, he explicitly endorses John Dupré’s “promiscuous realism” about natural kinds (Boyd 1999: 160; Dupré 1993). Natural kinds are relative to “disciplinary matrices” (scientific disciplines or groups thereof). And it is natural kinds that allow disciplinary matrices to achieve their “characteristic inductive, explanatory or practical aims” (Boyd 1999: 148). Disciplinary matrices are constituted in part by human interests, projects and practices and thus these interest, projects and practices are “partly definitive of natural kinds” (Boyd 1980: 642–3).



Picture 6.1 Kinds of kinds in Boyd and Kornblith

Boydian natural kinds are not confined to the disciplinary matrices of the natural sciences. Intellectual history and the social sciences have natural kinds, too. Thus feudal economy, capitalism, Islam, or Empiricism, are all natural kinds. Philosophy too has its natural kinds: thus disciplinary matrix, natural kind, and scientific knowledge are natural kinds in Boydian philosophy of science. Even individuals qualify as natural kinds. For Boyd there is no essential difference between natural kinds and natural individuals. Napoleon functions like a natural kind in political history (Boyd 1999: 154–6, 162–4).

Although Kornblith claims to be following Boyd’s understanding of natural kinds, his writings are not always clear on this point. For instance, in his 2002 book Kornblith works with two different conceptions of natural kind that I shall distinguish as the *wide* and *narrow* reading. On the wide reading, the opposite of a natural kind is a “*gerry-mandered kind*”, a kind held together by nothing but “our willingness to regard [it] as a kind” (2002: 10). A gerrymandered kind might consist of the tip of my nose, the Vienna *Hofburg* and the number 255. Boyd thinks of natural kinds as wide. The narrow conception of natural kind does not appear in Boyd: here natural kind contrasts with *social kind*. Both (narrow) kinds of kinds have theoretical unity, but whereas natural kinds are kinds introduced by natural science, social kinds are the kinds whose theoretical unity derives from a social role; presumably social kinds are the kinds of social science (Picture 6.1).

I mention this ambiguity in Kornblith’s talk of natural kinds, since at times he seems to me to conflate the two senses. Take for instance his claim that explanations based on social kinds are less deep than explanations in terms of natural kinds (2002: 10). That claim makes sense as long as we think of social kinds as gerrymandered. But it makes little sense if we think of the social kinds of “democracy” or “monarchy” within the disciplinary matrices of political theory or sociology. Or

consider Kornblith's contention that if knowledge were a social kind then epistemology would be less worthy of our attention (*ibid.*). It is hard to see why this would be so on the understanding of knowledge as a kind possessing theoretical unity. And yet, it has some (at least initial) plausibility on the rendering of knowledge as a gerrymandered kind.

Be this as it may, my main objection to Kornblith in this context is that he writes as if there were only *one single natural kind of knowledge*: knowledge is what cognitive ethology tells us it is. Coming from someone who claims to have adopted Boyd's theoretical framework, this is an odd view to take. After all, Boyd explicitly endorses a promiscuous realism about kinds. From a Boydian perspective, Kornblith's emphasis on cognitive ethology must be countered with questions like the following: Why isn't the kind "knowledge" used in the *sociology of knowledge* also a wide natural kind? Or why isn't the kind "knowledge" as used in our *folk psychology or folk epistemology* a wide natural kind, too? For the sociologist of knowledge, knowledge is "purely ... a natural phenomenon". Knowledge is "whatever people take to be knowledge". Knowledge consists of "those beliefs which people confidently hold and live by ... which are taken for granted as institutionalised, or invested with authority by groups of people". Knowledge is "what is collectively endorsed" (Bloor 1991: 5). This natural kind of knowledge does not coincide with the natural kind of knowledge investigated by cognitive ethologists. But this does not weaken its credentials within the disciplinary matrix of the social sciences.

A similar case can be made for our talk of knowledge in everyday life. It too qualifies as a natural kind for Boyd since a number of scientific disciplinary matrices (intellectual history for example) rely on this concept for its explanations and predictions. Note also that the everyday concept of knowledge differs from the concept of knowledge as it is used in cognitive ethology. In everyday life we attribute knowledge primarily to individuals, in cognitive ethology primarily to species.

How might Kornblith respond to the Boydian who advocates a promiscuous realism concerning the natural kinds of knowledge? Are there considerations that might be used to defend his idea that the cognitive-ethological natural kind of knowledge is the only one that really counts? Here are four possible answers.

According to the first proposal the cognitive-ethological kind explains the social role and thus the social kind of knowledge. Remember that for Kornblith the natural properties of gold explain its social role. I am not convinced. How does knowledge as species-wide reliable true belief explain (non-factive) knowledge as collectively sanctioned belief in a particular community? Even if it were to explain some parts of it, much would remain unaccounted for.

According to the second answer, cognitive ethology identifies the essential properties of knowledge. Other disciplinary matrices pick out merely its contingent or accidental properties. This too is not convincing. On the one hand, by what criterion are we to say that the ethologists' understanding offers the essential property? And, on the other hand, recall that the central property picked out by ethology—knowledge as reliably produced true belief—has been around in non-scientific armchair philosophy for a long time. That would not be good news for a naturalised epistemology taking its lead from the natural sciences rather than philosophy.

A third argument in defence of the primacy of the cognitive-ethological natural kind of knowledge is this: non-social animals can have knowledge, hence knowledge cannot be a social kind. This argument begs the question at issue. Why should we assume that the natural kind called knowledge by cognitive ethologists is the very same natural kind that sociologists or intellectual historians call knowledge? If they are not all the same kind, then the premise—non-social animals have knowledge—cannot establish the conclusion: knowledge is not essentially social.

The fourth, final, and ‘sledgehammer’ proposal is that only in the natural sciences do we have successful empirical theories. Fortunately Kornblith does not rely on this scientific idea.

Kornblith’s thesis that knowledge is a natural kind is directly linked to his claim that epistemologists have no reason to study our concept of knowledge or our intuitions regarding knowledge. In pressing this point Kornblith distinguishes two cases: that knowledge is a narrow natural kind like water, and that knowledge is a social kind (though a natural kind in the wider sense). His thesis is more plausible in the case that knowledge is a narrow natural kind. Chemists do not study water by studying our concepts of water (2007a: 39). Note however that it does not follow that our concepts (even our concepts of water) are without scientific interests. After all, concepts are—at least by Boyd’s reckoning—natural kinds too: they are central in the disciplinary matrices of cognitive psychology, linguistics, and the philosophy of mind.

Kornblith’s thesis—that concepts and intuitions are of little interest—is not very plausible, however, in the case of social kinds (2007b: 160–1). His dismissal of folk concepts and intuitions in the case of social kinds amounts to a rejection of “actors’ categories” as important to social science. This is a highly contentious claim. Can we really make sense of, say, “democracy”, “Islam” or “Empiricism” as wide natural kinds in political science or history, without paying attention to how these categories were understood by the historical actors themselves? Surely only in some pretty exceptional circumstances.

Let me sum up my difficulties with Kornblith’s claim that knowledge is a natural kind. First, Kornblith claims to follow Boyd, but his championing of the ethological natural kind of knowledge contradicts Boyd. Second, Kornblith does not offer a compelling argument for his choice of a privileged disciplinary matrix. Third, Kornblith’s dismissal of concepts and intuitions is problematic even if knowledge is a narrow natural kind. And fourth, it is even less convincing if knowledge is a social kind.

6.4 Craig’s Genealogy of Knowledge

Up to this point I have explained what Kornblith takes to be problematic about Craig’s genealogy of knowledge, and what I find unconvincing about Kornblith’s specific brand of naturalised epistemology. I now turn to Craig’s programme more directly. I first want to show that there is much more to Craig’s overall project than the one ten-line quotation that Kornblith focuses on. My goal here is not a detailed

explanation or reconstruction; my aim is a rough sketch of what I take to be a new interpretation of Craig's genealogy of knowledge.

Note first of all that—especially in his 1993 German-language book—Craig situates his project in the proximity to two somewhat unlikely bedfellows: Wittgenstein and natural science (1993: 37). Wittgenstein is an ally since he opposes conceptual analysis in terms of necessary and sufficient conditions, studies the function of concepts, and introduces the category of family-resemblance concepts. Craig's project has affinities with natural science in its method of hypothesis testing, the search for explanation, and a focus on evolution. Going beyond his wording, I would add *model-building* to the list: the building of simplified (and possibly even distorting) models of complex target systems.

Craig's model construction has two stages: the first focuses on the "epistemic state of nature", that is, a small community of language-using humans, engaging primarily in face-to-face communication, humans who are co-operative, dependent upon one another for information, and of unequal skills and talents. The central question regarding this state of nature is: Why would a concept like "knowledge" be introduced under these idealised—simplified and distorted—conditions? Craig answers that people in this situation have a salient need, to wit, the need to pick out and "flag" good informants. And the concept used to flag good informants is the core—or one central aspect—of our concept of knowledge.

As Craig emphasises more clearly in 2007 than in 1990 or 1993, this is not taken to be a historical thesis: the epistemic state of nature is not a historical period like the Pleistocene (2007: 191). It is rather an ubiquitous and important type of social-epistemic situation that one is likely to find in all human communities, past and present.

Craig goes to great length to show that his model of the epistemic state of nature passes the test of (what the philosophy of scientific models calls) "external validation". He does so by arguing that his model predicts and explains several of the features of our concept(s) of knowledge that have been identified in various philosophical theories. For instance,

- uses of "knowledge" without belief (Radford) (Craig 1990: 15–6)
- the role of counterfactuals (Nozick, Dretske) (Craig 1990: Ch. III)
- the role of causal relations (Goldman) (Craig 1990: Ch. IV)
- the role of methods (reliabilism) (Craig 1990: Ch. IV)
- the role of justifying reasons (internalism about justification) (Craig 1990: Ch. VIII)
- that all analyses have counterexamples (Gettier) (Craig 1990: Ch. VI) and
- the contextual variation in standards (Unger) (Craig 1990: Ch. XII).

These theories are often seen as excluding one another, but Craig thinks that his model can partially vindicate all of them: they contradict each other only if we over-generalise them, and only if we do not see that knowledge is a family-resemblance concept.

The second stage of Craig's model construction adds a *dynamic* dimension to the state-of-nature. The dynamic model takes the epistemic state of nature as its starting

point and seeks to track how the concept of knowledge would evolve and diversify as the simplifications and distortions of the state of nature are step by step reduced. This suggests that the dynamic model is really a form of “de-idealisation”. Craig focuses on a process he calls “objectivisation”: this is the process in and through which there emerges a variety of uses of “knowledge” that are no longer tied to face-to-face spoken communication, the needs of a specific hearer or questioner, short testimonial chains, and small communities. These new uses differ from the old one, amongst other things, in calling for higher epistemic standards for knowledge, and in breaking the link between knowledge and testimony. Craig stresses that objectivisation is not an ad-hoc stipulation, but a tendency that can be observed in the development of many concepts (1990: Ch. X).

The dynamic model too needs to pass muster as far as external validation is concerned. Craig suggests that it correctly predicts, or at least makes sense of,

- contexts with very high epistemic standards (1990: Ch. X),
- the distinction between know-how and know-that (1990: Ch. XVII),
- intuitions about lottery propositions (1990: XI), and
- our conflicting intuitions about epistemic scepticism (1990: XII–XIII).

The above is only a very rough indication of where I now see the originality of Craig’s project. In particular I hope to have made plausible that Craig’s project is in the proximity of naturalised epistemology: his explanatory strategy has more to do with model-building in the sciences than with the traditional search for necessary and sufficient conditions for concepts like “knowledge” or “justified belief”. It should also be obvious by now why I think Kornblith’s exclusive focus on the first few pages of Craig’s book is misleading.

6.5 Genealogy and Naturalized Epistemology

In the final step of my overall argument I shall try to relate Kornblith’s project and criticism to my interpretation of Craig’s programme.

I begin by restating what I have already emphasised at the end of my last section, to wit, that Craig’s affinities with natural science, but also the specific parallels with Wittgenstein, show that the intellectual distance between Kornblith and Craig is not as great as Kornblith alleges.

Note first of all that even though Kornblith regards the concept of knowledge as epistemologically uninteresting, what he does say about the concept fits nicely with the Wittgensteinian and Craigian emphasis on diversity and disunity. Kornblith writes that our folk concepts of knowledge “are not terribly unified ... [and] we will need to ... start presenting them in all their splendid disunity” (2007a: 43). I agree, though I would add that Craig’s epistemology offers one of the best means for capturing the disunity and unity in our concept of knowledge.

There is more of a tension between Kornblith and Craig at another point.

For Kornblith—as for Boyd—the central function of the concept of knowledge is to play a role in explanation and prediction. However, at least with regard to the first stage of his model—the static model of the epistemic state of nature—Craig rejects the focus on the need to explain others’ behaviour, and instead picks as central the need to flag good informants:

... the wish to explain, in some fashion, the behaviour of one’s fellows, ... (It ... has been suggested to me, that this idea could help us to see the concept of knowledge as some sort of theoretical construct, useful for explaining why other members of our community behave as they do.) But just how widespread this concern with explanation is ... is very hard to say ... it would not be advisable to allow ourselves such a starting point before we have exhausted the potential for far less contentious claims about the human situation ... (1990: 5)

This emphasis does mark a difference with Kornblith—but only as far as the first of Craig’s two models is concerned. Craig’s *dynamic* model (or at least a further development of it) leads to the prediction that further needs—in addition to the need to flag good informants—will also leave their marks on the concept. Consider two such needs: the need to explain behaviour, emphasised by Kornblith, and the need to mark the point at which we can terminate inquiry. That the latter is the main function of the concept of knowledge has been suggested by another critic of Craig, Klemens Kappel (2010). It is easy to see how these two needs become important in the process of objectivisation. Taking Smith to be a good informant regarding the location of tigers enables you to explain and predict some of his actions. There thus is a natural route from knowledge as a flag for a good informant to knowledge as essential to explanations of actions. Moreover, it is natural to shift from tagging the good informant to tagging the quality of the information he gives us. This involves a shift from “status” to “state”, a shift familiar from social-psychological investigations (cf. Barnes 2000). And thus there is a route also from flagging good informants to marking the point at which inquiry may cease.

But is there not a deep divide between Craig and Kornblith in so far as Craig speaks of knowledge as “... something that we delineate by operating with a concept which we create in answer to certain needs ...”? Does this not, as Kornblith alleges, really exclude the option that knowledge might be a natural kind? No, it does not. As Boyd emphasises: “... in a certain sense, human interests, projects and practices are partly definitive of natural kinds” (1980: 642). We build disciplinary matrices to satisfy certain of our needs and the explanatory, predictive and practical aims of disciplinary matrices determine which are the relevant concepts of natural kinds. This gives Craig all he needs to maintain his claim.

6.6 Conclusion

My aim in this paper has been twofold: to defend Craig’s programme against Kornblith’s criticism and to throw a critical light on Kornblith’s own proposal. I hope to have shown that the very heart of Kornblith’s naturalized epistemology—his treatment of natural kinds—is beset with problems, and that these problems

undermine his criticism of Craig. But this is not meant as a criticism of naturalized epistemology per se. It is merely to suggest that naturalized epistemology would profit from treating the genealogy of knowledge as a resource and ally rather than as an opponent. A concept-focused form of epistemology—at least Craig’s version of it—is not incompatible with naturalized epistemology; on the contrary the two approaches complement each other.

Acknowledgements For discussions relating to the topic of this paper I am grateful to Miranda Fricker, Hilary Kornblith, Martin Lenz, Veli Mitova, Adela Roszkowski, the students of my class on Craig in the Winter term 2009–2010, and audiences in Amsterdam and Berlin.

References

- Barnes, B. 2000. *Understanding agency: Social theory and responsible action*. London: Sage.
- Beebe, J.R. 2012. Social functions of knowledge attributions. In *Knowledge ascriptions*, ed. M. Gerken and J. Brown, 220–242. Oxford: Oxford University Press.
- Bloor, D. 1991. *Knowledge and social imagery*, 2nd ed. Chicago: University of Chicago Press.
- Boyd, R. 1980. Scientific realism and naturalistic epistemology. In *PSA: Proceedings of the biennial meeting of the Philosophy of Science Association, Vol. 2: Symposia and invited papers*, 613–662.
- Boyd, R. 1983. On the current status of the issue of scientific realism. *Erkenntnis* 19: 45–90.
- Boyd, R. 1988. How to be a moral realist. In *Essays on moral realism*, ed. G. Sayre-McCord, 181–228. Ithaca/London: Cornell University Press.
- Boyd, R. 1991. Anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies* 61: 127–148.
- Boyd, R. 1999. Homeostasis, species, and higher taxa. In *Species: New interdisciplinary essays*, ed. R.A. Wilson, 141–185. Cambridge, MA: MIT Press.
- Craig, E. 1990. *Knowledge and the state of nature: An essay in conceptual synthesis*. Oxford: Clarendon.
- Craig, E. 1993. *Was wir wissen können*. Frankfurt a. M: Suhrkamp.
- Craig, E. 2007. Genealogies and the state of nature. In *Bernard Williams*, ed. A. Thomas, 181–200. Cambridge: Cambridge University Press.
- Dupré, J. 1993. *The disorder of things: Metaphysical foundations of the disunity of science*. Oxford: Oxford University Press.
- Fricker, M. 2007. *Epistemic injustice: Power and the ethics of knowing*. Oxford: Oxford University Press.
- Fricker, M. 2010. Scepticism and the genealogy of knowledge: Situating epistemology in time. In *Social epistemology*, ed. A. Haddock, A. Millar, and D. Pritchard, 51–68. Oxford: Oxford University Press.
- Gelfert, A. 2011. Steps to an ecology of knowledge: Continuity and change in the genealogy of knowledge. *Episteme* 8: 67–82.
- Hacking, I. 2007. Natural kinds, rosy dawn, scholastic twilight. *Royal Institute of Philosophy Supplement* 82(61): 203–239.
- Kappel, K. 2010. On saying that someone knows: Themes from Craig. In *Social epistemology*, ed. A. Haddock, A. Millar, and D. Pritchard, 69–88. Oxford: Oxford University Press.
- Kelp, C. 2011. What’s the point of ‘Knowledge’ anyway? *Episteme* 8: 53–66.
- Kornblith, H. 1993. *Inductive inference and its natural ground: An essay in naturalistic epistemology*. Cambridge, MA: MIT Press.
- Kornblith, H. 2002. *Knowledge and its place in nature*. Oxford: Clarendon.

- Kornblith, H. 2006. Appeals to intuition and the ambitions of epistemology. In *Epistemology futures*, ed. S. Hetherington, 10–25. Oxford: Oxford University Press.
- Kornblith, H. 2007a. Naturalism and intuitions. *Grazer Philosophische Studien* 74: 27–49.
- Kornblith, H. 2007b. The metaphysical status of knowledge. *Philosophical Issues* 17: 145–164.
- Kornblith, H. 2011. What should we care about the concept of knowledge? *Episteme* 8: 38–52.
- Kusch, M. 2009. Testimony and the value of knowledge. In *Epistemic value*, ed. A. Haddock, A. Millar, and D. Pritchard, 60–94. Oxford: Oxford University Press.
- Kusch, M. 2011. Knowledge and certainties in the epistemic state of nature. *Episteme* 8: 6–23.
- Neta, R. 2006. Epistemology factualized: New contractarian foundations for epistemology. *Synthese* 150: 247–280.
- Pritchard, D. 2009. *Knowledge*. Houndmills: Palgrave MacMillan.
- Williams, B. 2002. *Truth and truthfulness: An essay in genealogy*. Princeton/Oxford: Oxford University Press.

Part II
**Shaping the Norms of Our Intellectual and
Practical Engagement with the World**

Chapter 7

Sensibility and Metaphysics: Diderot, Hume, Baumgarten, and Herder

Stephen Gaukroger

7.1 Introduction

Reconciling, or choosing between, the competing demands of reason and sensibility was the most urgent philosophical concern of the second half of the eighteenth century. In pursuing this question today, I want to begin by looking at two developments that, when combined, prepare the ground for the general problem of reason and sensibility that I'll be focusing on. The first of these is the generalization of what in Locke is a dichotomy between reason and sensation into one of reason versus sensibility. This effectively begins with French Lockeans such as Condillac, and is continued in Buffon, and Diderot, on whom I will focus. What at first looks like something narrowly epistemological turns out to have a central moral and social dimension. The second development is the radicalization of the dichotomy in Hume, who stresses both the inadequacy of reason, but also, unlike Rousseau for example, the necessity of reason. This generates a dilemma. Combined with the first development, which highlights just how important it is to get the relation between reason and sensibility right, the Humean dilemma takes on a wholly new significance. I shall look at two opposed solutions to the dilemma, on the question of how the standing of reason might be saved. The first is that of Baumgarten, who tries to rationalize sensibility by showing how it can be subordinated to reason. The second is that of Herder, who takes an already-naturalized sensibility and makes reason commensurate with it through a process of naturalization.

S. Gaukroger (✉)
History of Philosophy and History of Science,
University of Sydney, Sydney, NSW, Australia
e-mail: stephen.gaukroger@sydney.edu.au

7.2 Diderot

Questions of the connection between sensibility and sensation were introduced into French thought in the 1740s in the work of Condillac and Buffon. Both independently used the image of human statue which is progressively given sensory faculties, in order to explore just what is needed to produce a conscious self-aware human being.¹ We are invited to imagine a human statue which we bring alive by attributing, one by one, various sensory faculties to it, asking what its experience of the world would be like. Imagine we give the statue the power to smell and place a rose in front of its nose, so that the statue experiences the odour of a rose. There is nothing in this experience that would lead the statue to imagine that that experience had an external source, so the olfactory sensation is not experienced as being that of an external object, but simply as an experiential state. Similarly with the power to hear, to see, to taste. Even a statue with all five of the senses, which was able to compare, reflect, remember, and accomplish the other intellectual operations, would not be led, on these grounds alone, to imagine that its states were anything but internal and self-contained. Despite having a grasp of spatial relations, for example, the statue would be solipsistic: its sensory experiences would not project it into, or connect it with, the world. On the contrary, it would remain isolated and self-contained. The question now is how it is possible for us to develop a conception of the world as independent of us. Buffon's view was that only touch can provide a sense of an external world, and with this sense effectively comes a precondition for morality, in that with it comes a recognition of the existence of others, and a love for them replaces the earlier narcissism. Condillac's answer to the question is that a sense of reflection emerges from the sensations, allowing us to distinguish our own body from the sensations themselves, by combining the sensations and making something new out of them, in which the various objects of sensation can be compared under different descriptions. The crucial point is that, for both Buffon and Condillac, sensibility and a recognition of an external world come simultaneously.

For those thinkers in the French Lockean tradition, it is axiomatic to the sensationist project that one begins life with a *tabula rasa*, and the question is how one develops a cognitive, affective, and moral life on this basis. Diderot develops this line of thought most fully, explicitly linking sensation with moral and aesthetic sensibility. He does this through an exploration of the 'mentality' of someone who is deprived of sensory capacities in some way. In his account of blindness, for example, he compares the blind with the sighted in order to explore what this tells us about sensibility in general.² A deficient sensibility is primarily a question of an emotional, aesthetic, and moral challenge for Diderot, and because of their impoverished sensibilities, the blind turn their minds inwards and are drawn to thinking in terms of abstractions. The blind offer a crucial case study for Diderot because he believes

¹ Condillac (1754), Part I; Buffon, in the section "Des senses en general" in *De l'homme*: Buffon George Louis Leclerc (1827, xii: 165–86).

² Diderot (1749).

that their abstract manner of experiencing pain in others weakens their sense of sympathy for the suffering of others. His target is not confined to the blind, however. The blind just manifest in a particularly concrete way a general insensibility to the world on the part of those who think and experience it in terms of abstractions. He is particularly concerned with Cartesians in this French context, but the German metaphysical tradition, as represented by Wolff, would be a prime contender for this status also. For Diderot, our relation to the world depends very much upon how we arrive at that relation, and one assimilates cognitive information in a process which is always and necessarily social, cultural, and has moral implications, so that what is shaped is not merely a cognitive sensibility but a sensibility in which cognitive, affective, and moral questions are inextricably tied together. What is ultimately at stake is the sensory basis of civic life, where the contrast is between sensibility and solipsistic rationalism.³ The general question underlying this is that of where the ideas that regulate our lives – our moral, emotional, social, political and intellectual lives – come from. The approach of those metaphysicians and others who conceive of the world in abstract terms, and as a consequence examine our relation to it in these same terms, now becomes not merely misguided but socially and morally irresponsible.

7.3 Hume

By the middle of the eighteenth century, the question of the relation between reason and sensibility had started to become a pressing problem, and it was Hume's formulation, more than any other, that was to challenge philosophers, particularly in Germany, from the 1760s onwards. Hume raised a dilemma with respect to the nature of understanding. Simplifying somewhat, he came to the view that the most developed form of systematic general understanding, metaphysics, led to paradoxical conclusions which were so contrary to "common life" that we could not accept them as true. Yet to the extent to which we desire to develop a critical understanding of the world and our moral and other practices, we cannot renounce such forms of philosophical enquiry. We are faced with a mode of life and thought that leads to dangerous or ridiculous phantasies, but which is part of the thoughtful existence distinctive of civilization, by contrast with an unreflective mode which he associated with barbarism.

We can think of Hume as generalizing from an account of natural philosophy developed by Locke. For Locke, there is a shifting balance between the demands of systematic explanation and those of experimental natural philosophy. Hume in effect transfers this view to knowledge in general. He does not deny that understanding can take a systematic form, only that the kind of understanding that can take a systematic form – and he is concerned with metaphysics in this context – cannot

³ See the account in Riskin (2002, ch. 2).

be constitutive of our understanding of the world: it can be part of it (and indeed, he believes, by contrast with his later “common sense” school critics such as Reid, *must* be part of it), but not all of it. Hume offers an account of “human understanding” whereby it is a judicious balancing of propositional and non-propositional considerations, of considerations of reason and sensibility.

Hume’s use of sceptical arguments to undermine the claims of systematic metaphysics – he takes Malebranchian metaphysics as its most advanced form – came in time to be immensely influential, but equally important was his own practice as an essayist and historian, which had a more immediate impact. It was in this latter capacity that his life and work manifested the other side of the balance between systematic and non-systematic enquiry. The essays, in particular, balanced the sceptically-driven metaphysical writings, and the two represented two different *personae*, both equally necessary. Here we have what is most distinctive about Hume’s position, namely his refusal to choose between systematic and non-systematic considerations. His combination of a commitment to a sceptically-driven form of enquiry as absolutely central to our critical engagement with the world in which we live, on the one hand, and his demonstration of the absurdities to which this leads, on the other, leads Hume to a distinctive view of the role and nature of judgement. His view is that understanding consists in the exercise of judgement between considerations that are not exclusively propositional.

What is distinctive about this judgement is that it is not a “philosophical” judgement as such. That is to say, it is not a judgement that the philosopher, *qua* philosopher, is best qualified to make. If one wants to judge the “good life,” this is not something that any particular group – theologians or metaphysicians – have any real claim on. But as traditionally conceived, for example if one thinks of the philosophical tradition from Plato and Aristotle, an answer to the question how to live is precisely the ultimate goal of philosophy, and it is a goal that establishes its autonomy from theology for example. Even worse, metaphysics and other disciplines are compared in terms of what Hume terms “natural history,” something deriving from the experimental philosophy tradition and its empiricist successor, and which he uses to naturalize enquiry on the one hand, while denying either systematicity or even a unified programme to natural philosophy on the other.

7.4 Baumgarten

It did not need Hume’s attack on metaphysics to put its aspirations in question. Systematic metaphysics did not fare at all well in the early to mid eighteenth century, and was regaled across the spectrum from radical Parisian *philosophes* to conservative Jesuits. The only metaphysical programme of any significance was that of Wolff, for whom it was part of the reform of the German university system.⁴

⁴ See Schneiders (1985a), Hinske (1983), Bödeker (1990).

Distinguishing between “historical” knowledge, which concerns itself with empirical and causal matters, and “philosophical” knowledge, which works from basic principles, Wolff rejected the traditional ascendancy of theology, jurisprudence, and medicine, arguing that they were merely “historical” forms of enquiry, whereas philosophy could raise knowledge to the level of what the scholastics had termed *scientia*: comprehensive understanding based on first principles. Wolff was fighting a losing battle here,⁵ but the crucial thing is that the reason/sensibility distinction gets translated into a specific metaphysics versus “empirical history” contrast in the German context.

What Wolff offered was a basically scholastic metaphysics reformulated along Leibnizian lines. Wolff began by rewriting logic in Leibnizian terms, conceiving its aim to be the provision of a means whereby confused concepts were analyzed and made clear and distinct. This procedure was then applied to the whole of philosophy, conceived as a discipline that investigated why things are as they are. The model was explicitly that of mathematics, where necessary truths were revealed by establishing the systematic connections that propositions have with one another. The systematic metaphysics that results on this conception has no place for sensation, which it construes simply as a form of cognition, namely sensory perception, which produces only obscure ideas.⁶

But by mid-century, such a dismissive view of sensibility was simply not viable, and the central problem for those pursuing philosophy via metaphysics was what account to give of sensibility. Wolff had made it clear that every area of enquiry should be pursued with scientific rigour, but he had done little to further the study of the arts and humanities on this basis. The one significant development of the Wolffian programme, in Baumgarten’s *Aesthetica* of 1750, was on this question, and what he proposed was a science of the “lower faculties.” He had already noted in his early *Meditationes* that, in promising to improve our knowledge, logic actually promises more than it can deliver, for logic deals with intellectual cognition, whereas human cognition generally also involves sensory cognition, so if we rely on logic alone, we will leave a whole realm in its original unanalyzed state.⁷ In particular, if we allow the intellect to rule the lower cognitive faculties, this will have detrimental consequences for art, producing something frigid and scholastic, for we find a richness and vividness in sensory perception that is lacking in purely intellectual cognition.⁸ Baumgarten terms the study of sense perception “aesthetics,” and he argues that it is the “younger sister” of logic: whereas logic deals with intellectual cognition, aesthetics deals with the “lower cognitive powers,” namely sensory cognition. Indeed, logic is clearly the model: the role of logic on this conception is to make confused concepts distinct through analysis, and Baumgarten’s aim is to

⁵ See Hammerstein (1983a, b).

⁶ See, for example, Wolff (1732, §38).

⁷ Baumgarten (1735, §115).

⁸ Baumgarten (1750–1758, §§ 105, 560, 619).

perfect our powers of perception.⁹ As a means to understanding what this would consist in, he sets out to establish that perfected perception is equivalent to beauty, so that his “science of sensory cognition” is also a “critique of taste.”¹⁰ That is to say, he makes a direct link between sense perception and aesthetic sensibility. Moreover, there is a set of rules by which the perfectibility of the two can be achieved. One important difference between intellect and sensory faculties is that whereas the former operates by means of demonstrative inference, the confused elements of the latter are related through association, and Baumgarten identifies “the law of the imagination”, namely, that ‘the partial perception of an idea conjures up the whole idea,’ with the law of association of ideas.¹¹ Just as there is a “natural logic,” that is, a grasp of inferential relations prior to any education in logic, so too is there a “natural aesthetic” which children exercise in their sensory exploration of the world, and above all in playing games.¹² And just as natural logic can be perfected through training in rules developed through an analysis of inference, so too can natural aesthetics be improved in an analogous way.

The model for analyzing the sensory realm is drawn exclusively from that devised for analyzing the intellectual realm, and aesthetics falls under a general comprehensive metaphysics for Baumgarten. The fact that it works via association of ideas rather than deductive inferences between ideas does not mean that aesthetics qua the study of sensation is a genuinely autonomous discipline. Quite the contrary, it remains resolutely subservient to reason for Baumgarten. Here, then, we have a model of how we might take sensibility seriously, yet subsume it under reason. But it was not a model that was taken up in the middle decades of the eighteenth century.

In France and Germany, the question of the relation between sensibility and reason became highly polemicized in the 1750s and 1760s through the influence of Rousseau. In Britain, Rousseau’s influence was significantly weaker, but what Rousseau represented for many philosophers – an essayist of wit and urbanity whose philosophical interests went well beyond those subjects that fell under traditional metaphysics, and who represented a philosophical *persona* quite different from the metaphysician – was something that could also be found in Hume. Moreover, both dealt with moral, political, educational, and pedagogic questions – those questions that would fall under the “moral sciences” – in terms of anthropology rather than either metaphysics, or Christian teaching, or civic humanism (as traditionally conceived). Although Rousseau and Hume drew different conclusions on many of the questions they explored, what was crucial was their way of pursuing the subject, and

⁹ Baumgarten (1750–1758, §§ 1–13). Cf. Baumgarten (1739, §§ 519–33). There are a number of developments in Baumgarten’s thinking between 1739 and 1750, but the differences need not concern us here, and in any case are slighter than was once thought: see Franke (1972). Cf. École (1980).

¹⁰ Baumgarten (1739, §§ 521, 622).

¹¹ Baumgarten (1750–1758, §561).

¹² Baumgarten (1750–1758, §§ 54–5).

their questioning of the power of reason to regulate a range of issues from morality and aesthetics to epistemology and religious belief.

The two tend to be drawn together in Germany, where we find the influence of Rousseau balanced to some extent by Hume, especially inasmuch as Hume is associated with the project of a naturalized psychology. In mid-century, British and French ideas were introduced into Berlin, in particular, in a context of growing hostility to Wolffian *Schulphilosophie* and a revival of what was seen as Thomasius' eclecticism.¹³ Above all, the material from France and Britain was disseminated through journals announcing their subject matter as *schönen Künste und Wissenschaften*. "Beautiful science" or "beautiful arts" correspond to the French term *belles lettres*: humane learning, by contrast with logic, metaphysics, and theology. What is happening in this literature is that there is a shift of learning generally, including the sciences, from an academic culture into an essay culture, one whose audience is the Republic of Letters.¹⁴ Although the shift might be viewed as bringing Germany into alignment with Britain and France, the relative suddenness and abruptness of the transition, together with the fact that the leading figures – Nicolai, Lessing, and Mendelssohn – were incorporating this shift into an explicit programme of cultural and political reform, meant that the rejection of metaphysics, as a vehicle for thought about and engagement with fundamental issues about the nature of the world and our place in it, raised the whole question of the value of particular modes of enquiry, and in particular whether any overarching form of understanding was possible.

7.5 Sensibility

Hume had rejected the idea that there is some meta-discipline, traditionally identified with metaphysics, that can stand over all other disciplines and critically evaluate them, without itself being subject to any such evaluation. But he did employ an overarching form of enquiry himself, which he refers to as "natural history." Somehow, it plays this role without being immune to criticism itself, and without being uniquely empowered to offer criticism of other disciplines.

What exactly is this "natural history?" Natural history in the literal sense is a classificatory discipline, traditionally associated with botany, zoology, and mineralogy, but it acquires more complex senses in Bacon at the beginning of the seventeenth century, and then in Boyle and the defence of "experimental natural philosophy" in the second half of the seventeenth century. The central contrast for the "experimental natural philosophy" tradition is between, on the one hand, an experimental form of natural philosophy which identifies and accounts for problematic phenomena at a purely phenomenal level without recourse to any idea of an underlying causal level; and,

¹³ See Schneiders (1985b) and Holzhey (1983: 19–29).

¹⁴ Schneiders (1983) and Zammito (2002, 8–13).

on the other hand, systematic natural philosophy, which refers explanations to an underlying microscopic level, on the understanding that this is where the real action happens. What is at issue is whether all explanations must take the form of deducing the macroscopic behaviour of bodies from the behaviour of their microscopic constituents. Advocates of experimental natural philosophy – beginning with Boyle in his pneumatics and Newton in his account of spectral colours, and effectively dominating chemistry and the study of electricity by mid-century – urged the legitimacy, and fruitfulness, of drawing connections between phenomena. If we take natural history in this sense, then we can perhaps see how we might conceive of a discipline that can in some way help us to make sense of the relation between reason and sensibility, without any pretensions to understanding these matters in terms of a system resting on a small number basic truths, and without the aspiration to offer some absolute critical perspective, itself immune to criticism.

It is, I think, a natural assumption that what we need to work on if we are to achieve this goal is reason. Somehow, reason needs to be reformed if it is to abandon its aspiration to play the role of the ultimate critic, while itself remaining immune from any form of criticism. It is of course its relation with sensibility that presents the problems, so sensibility has to be understood properly, but this could presumably be achieved without reason shedding its autonomy. Another approach, a very different one, is to examine how one accounts for sensibility and then to ask whether we might approach reason from the perspective of sensibility.

One apparent obstacle to this strategy is the sheer diversity of the phenomena that fall under the rubric of sensibility, especially when compared to reason, which gives every appearance of being intrinsically unified. But the diversity is not in fact as much of an impediment as it first appears. We can distinguish four phenomena that fall under the term, as used at the time: aesthetic sensibility, a kind of awareness or appreciation that accompanies perceptual cognition in varying degrees; moral sensibility, an ability to exercise untutored moral discrimination; sensation, something that figures in our cognitive relation to the world; and sensitivity, something that accompanies certain kinds of physiological activity. These are not unconnected phenomena. As Anne Vila has pointed out, “sensibility was the essential link between the human body and the psychological, intellectual, and ethical faculties:” it helped unify the human faculties, as it was “seen as the root of all human perceptions and reflections, as the innate and active principle of sociability that gave rise to human society, as a kind of sixth sense whose special affective energy was essential to both virtue and to art, and finally, as the paradigmatic vital force whose actions could be detected in every bodily function, be it healthful or morbid.”¹⁵

One crucial unifying feature of the diverse phenomena that came under the rubric of sensibility lay in how they were being accounted for. One of the distinctive features of sensibility is the extent to which it was becoming naturalized by the middle of the eighteenth century. This took many forms, from the attempts of moral philosophers such as Hume and others to treat political, social, and economic questions

¹⁵ Vila (1998, 2).

in terms that went beyond both traditional Christian thinking and civic humanism, to the attempts of those working within the life sciences, most notably Buffon, to extend their account of zoology into the human and social realm. Its characteristic feature is the replacement of speculative by empirical forms of enquiry. This has consequences for what we consider the role of reason to be, and how we think of reason. The key question is that of how our beliefs and values are shaped. Whereas philosophers had traditionally assumed that reason and the passions played the crucial role here, typically (if not universally) thinking of the passions as an obstacle to reason, we are led in a different direction if we approach reason from a consideration of sensibility.

7.6 Herder

This is the direction that Herder takes. Reason is not simply assimilated to sensibility. It becomes subjected to an empirical, comparative form of treatment which actually emphasizes its distinctive characteristics.

Wolff had distinguished “rational” from “empirical” psychology, and the latter plays a key role in Herder’s strategy. It worked on analysis of the behaviour of others and on analysis of our own psychological states, and he writes that it “is really a *history* of the soul and can be known without any other discipline.”¹⁶ Empirical psychology provided the raw materials for rational psychology, in that the task of rational psychology was to derive the results of empirical psychology from a priori first principles, with a view to achieving clarity and distinctness. This kind of move from empirical to rational psychology was a form of subsumption of the former under the latter. At the same time, however, for Wolff and Baumgarten there is a continuum between the higher and lower faculties, between understanding and sensibility.¹⁷ This opens up the possibility of another form of enquiry, one in which the aim is to explore the relation between the faculties as placed along the continuum. This latter would be a form of “history” in Wolff’s terms, “natural history” in Hume’s terms. In itself, it would not rule out the project of subsuming sensibility under reason, but if one had independent objections to such a project – objections which were plentiful from the 1750s – then it would provide a potentially promising route. What, in Wolffian terms, were the lower faculties would now be reassessed, being called upon to provide the model for analysis of the higher ones. It was a matter of using the resources employed to explore sensibility, resources that were highly naturalized, to explore the understanding.

The enquiry now becomes empirical, not speculative, quite different from the facultative logic that the Wolffian tradition takes over from Leibniz as the route to clarification of ideas. I want to distinguish two sets of resources that help Herder to

¹⁶ Wolff (1996, §34).

¹⁷ See the discussion of these questions in Zammito (2002, ch. 2).

follow the empirical path. The first is something that he takes over from developments in the life sciences, and more broadly in matter theory. It offers a picture of natural philosophy which makes possible an understanding of living things, which the earlier mechanist model had unsuccessfully tried to account for in reductive terms. A new theory of matter construes it as active, not inert, and it introduces developmental considerations into natural history in a fundamental way. The second is an extension of naturalizing resources to include language, so that it is no longer merely a means of expression of thoughts, but something that goes to the heart of what it is to think. Analyzing thought through language offers a potential avenue for naturalization not available if language is simply a means of expression which has no content in its own right.

Let us consider the natural-philosophical question first. Sensibility could hardly act as a model of naturalization if it itself could not be naturalized. If sensation and sensitivity could not be accounted for in naturalistic terms, for example, then the project clearly could not get off the ground. The mechanist model of natural philosophy, whose most successful version was that offered by Descartes, was premised on the notion of inert matter: the underlying idea was that the mechanically-characterized behaviour of inert microscopic corpuscles of matter was constitutive of all physical phenomena. On this conception, such phenomena as nervous sensibility were reduced to biomechanics – basically, inert matter moving under pressure. By the early decades of the eighteenth century, the limits of this form of reductive explanation became evident in chemistry, electricity, and physiology, as well as in the life sciences. In its place, there developed a notion of matter as inherently active, with electrical conductivity, chemical reactions, and muscular and nervous action as paradigmatic examples of this activity. The most striking example, however, came from the life sciences, notably from the discovery, in 1740, of the reproductive behaviour of the freshwater polyp, or hydra. The hydra regenerated by budding, like a plant, but even more remarkably, when cut into multiple parts, each part would regenerate, with all the features of the whole polyp – mouth, arms, legs, stomach – appearing, no matter from what segments of the original the cut part came. Matter, it seemed, contained vital powers. Matter as envisaged by mechanists was not matter *per se*, as they believed, but simply dead matter: that is, matter that had lost all its interesting active and vital powers, and was reduced to a mere shell. Indeed, advocates of the new understanding of matter, such as Maupertuis, replaced the mechanist micro-/macro- distinction with an inner/outer one. It was not the smaller parts of matter that explained the behaviour of the larger ones, but the inner properties of matter that explained the outer behaviour: size was replaced by depth. It was not just vital phenomena that fared better on this conception, but a whole range of things that were wholly inexplicable on the corpuscularian model: why chemical substances reacted in constant proportions, and why no mechanical force could separate the original components once they had combined; why in physiological processes such as reflex action, the response far exceeded the stimulus in force, in apparent contravention of the basic laws of mechanics; and so on.

As well as a new understanding of matter as active, allowing the vital phenomena that made up sensibility to be incorporated into a general form of natural philosophy, there was a second development that had a more direct impact on Herder's project.

Maupertuis, in studying heredity, had been led to an appreciation of developmental history, and Buffon had argued for the indispensability of a developmental perspective in understanding natural history. Indeed, Buffon took developmental questions as providing the explanatory core of natural history, engaging the methodological questions of truth and evidence that this raises, above all the question of how, using only what we can observe at present, we can reconstruct the past of the solar system and the earth, for example. At the same time, in the extensive Preliminary Discourse to the *Encyclopédie*, d'Alembert and Diderot were accounting for the present state of knowledge in terms of a genealogy of reason, showing how religious and other obstacles to learning had not only hindered but skewed scientific understanding, in the process identifying just what it was that had been necessary for the present state of scientific understanding to be reached. Meanwhile, Hume's *History of England* and William Robertson's *History of Scotland* made comprehensive attempts to deal with questions in political philosophy through detailed historical analysis.

There are, then, two mid-century developments on which Herder can rely. The first was a confidence that, with the rethinking of matter as something active, basic phenomena of sensibility can be dealt with at a reasonably fundamental level as part of a comprehensive natural philosophy. The second was the existence of models of explanation through developmental history.

We can think of Herder's project in terms of a threefold naturalization. He approaches reason by understanding it in terms of thought; he approaches thought via what the Wolffian tradition characterized as empirical psychology; and the tool that he uses in this empirical psychology is language, which he treats as an empirically analyzable manifestation of thought. Moreover, it is an empirically analyzable manifestation that has a historical dimension: the history of languages – "history" in the eighteenth-century sense, as something that has both a diachronic and a synchronic dimension – is a history of thought, which in turn is a history of reason, albeit one of a far more subtle form than the "triumph of the present" type history of thought offered by Diderot and d'Alembert.

The move from reason to thought is a move from speculative metaphysics to empirical psychology. In itself this was not especially contentious, and empirical psychology played a crucial role in the Wolffian tradition. But this role was above all a necessary preparatory one, organizing the raw materials upon which an a priori structure could then be imposed, and thereby yielding a degree of clarity and understanding which empirical psychology in its own right could not aspire to. The first stage was as far as the Wolffian programme could get with sensibility, at least before its reform by Baumgarten, so that it was considered that the problem with sensibility was that it could never aspire to anything beyond the kind of limited and necessarily obscure comprehension offered by empirical psychology. But the project of providing metaphysical structure had now fallen into disrepute outside what was by this stage a very small Wolffian circle, not least because of Hume's onslaught. In its place, what was now effectively being proposed was that the kind of understanding offered by empirical psychology was really all we had as far as understanding was concerned, and that with an expansion of its resources, it was actually a much more powerful explanatory tool than its treatment at the hands of metaphysicians would suggest.

If empirical psychology was to go beyond the role it played in the metaphysical project, if it was to offer an analysis of thought that had genuine aspirations to clarity and understanding, thought, in all its distinctiveness, had to be captured somehow. Reason could not simply be rejected in favour of sensibility, along Rousseauian lines, nor could thought simply be reduced to a form of physiological activity, along the lines that La Mettrie was proposing. If the project was to have any plausibility, thought had to retain some autonomy: there had to remain something distinguishing about it which successful explanation was able to capture. Language and thought had been associated from antiquity, and Descartes had argued in his *Discours de la méthode* that the fact that human beings have language but animals have none “shows not only that animals have less reason than men, but that they have none at all.”¹⁸ The close relation between language and rational thought was a given of seventeenth- and eighteenth-century culture. But Herder wants and needs something stronger than this: he wants the analysis of language to be an analysis of thought.

In his *Fragments on Recent German Literature*, from 1767 to 1768, he writes that language is “a form of cognition not merely in which, but also in accordance with which, thoughts take shape, where in all parts of literature thought is tied to/follows/adheres to [*klebt*] expression, and forms itself in accordance with the latter.”¹⁹ Such linguistic naturalism has an immediate advantage over materialist forms of reductionism, for these latter were widely deemed to commit one to determinism, whereas the linguistic naturalism that Herder is offering here clearly has no such consequences. It simply has no bearing on the question of determinism. Linguistic naturalism also has a positive advantage in that it allows a comparative form of study. By comparing earlier languages with present ones, and by comparing contemporary languages with one another, one is comparing forms of thought. Seventeenth- and eighteenth-century accounts of morality had had great problems in dealing with moral diversity. There was a concerted attempt in Locke, Shaftesbury, and Hutcheson for example, to establish some universality in morality. But it is not so much the failure of this project that is interesting, so much as the inability of prevailing theories of moral sensibility to take into account, in a theoretical way, the fact of widespread moral diversity: at least outside those accounts that simply traced it back to the Fall, where they postulated the emergence of a kind of moral Tower of Babel. I characterized moral sensibility as a form of sensibility earlier, in setting out the understanding of what the term included in the *Encyclopédists* and others, but simply treating it as coming under sensibility *as opposed to reason* is going to raise fundamental problems about moral responsibility. It is particularly apparent in social reductionist accounts of morality such as that of Helvétius. To the extent to which thought can be incorporated into, but not reduced to, the naturalist schema, Herder’s approach has a potentially great advantage. For linguistic variability is ultimately a variability in what would come to be termed mentalities, and moral variability is a form of divergence in mentalities, something that might still be

¹⁸Descartes (1974–1986, vi: 58).

¹⁹Herder (2002, 48) (translation altered slightly).

considered problematic and indeed worrying, but no longer treated as bizarre and incomprehensible.

The transformation of the study of reason into an empirical discipline inaugurates a form of anthropology which was in direct competition with metaphysics.

References

- Baumgarten, Alexander Gottlieb. 1735. *Meditationes philosophicae de nonnullis ad poema pertinentibus*. Halle: Grvnertus.
- Baumgarten, Alexander Gottlieb. 1739. *Metaphysica*. Halle: Renger.
- Baumgarten, Alexander Gottlieb. 1750–1758. *Aesthetica*. 2 vols. Frankfurt an der Oder: Baudouin Frères and N. Delangle.
- Bödeker, Hans Erich. 1990. Von der “Magd der Theologie” zur “Leitwissenschaft”: Vorüberlegungen zu einer Geschichte der Philosophie des 18. Jahrhunderts. *Das achtzehnte Jahrhundert* 14: 19–57.
- Buffon George Louis Leclerc, Comte de. 1827. *Œuvres complete*. 34 vols. Paris: Delangle Baudouin.
- Condillac, Etienne Bonnet de. 1754. *Traité des sensations*. Paris: de Bure.
- Descartes, René. 1974–1786. *Oeuvres de Descartes*, ed. C. Adam, and P. Tannery. 11 vols. Paris: Vrin.
- Diderot, Denis. 1749. *Lettre sur les aveugles, à l’usage de ceux qui voient*. London: Arnold.
- École, Jean. 1980. *La Métaphysique de Christian Wolff*. 2 vols. Hildesheim: Olms.
- Franke, Ursula. 1972. *Kunst als Erkenntnis: die Rolle d. Sinnlichkeit in d. Ästhetik d. Alexander Gottlieb Baumgarten*. Wiebaden: Steiner.
- Hammerstein, Notker. 1983a. Die deutschen Universitäten im Zeitalter der Aufklärung. *Zeitschrift für historische Forschung* 19: 73–89.
- Hammerstein, Notker. 1983b. Christian Wolff und die Universitäten. In *Christian Wolff 1679–1754*, ed. Werner Schneiders, 266–277. Hamburg: Meiner.
- Herder, Johann Gottfried. 2002. *Philosophical writings*. ed. and Trans. Michael N. Forster. Cambridge: Cambridge University Press.
- Hinske, Norbert. 1983. Wolffs Stellung in der deutschen Aufklärung. In *Christian Wolff 1679–1754*, ed. Werner Schneiders, 306–320. Hamburg: Meiner.
- Holzhey, Helmut. 1983. Philosophie als Eklektik. *Studia Leibnitiana* 15: 19–29.
- Riskin, Jessica. 2002. *Science in the age of sensibility: The sentimental empiricists of the French enlightenment*. Chicago: Chicago University Press.
- Schneiders, Werner. 1983. Zwischen Welt und Weisheit. Zur Verweltlichung der Philosophie in der frühen Moderne. *Studia Leibnitiana* 15: 2–18.
- Schneiders, Werner. 1985a. Der Philosophiebegriff des philosophischen Zeitalters: Wandlungen im Selbstverständnis der Philosophie von Leibniz bis Kant. In *Wissenschaft im Zeitalter der Aufklärung*, ed. Rudolf Vierhaus, 58–92. Göttingen: Vandenhoeck & Ruprecht.
- Schneiders, Werner. 1985b. Vernünftiger Zweifel und wahr Eklektik: Zur Entstehung des modern Kritikbegriffs. *Studia Leibnitiana* 17: 143–161.
- Vila, Anne C. 1998. *Enlightenment and pathology: Sensibility in the literature and medicine of eighteenth-century France*. Baltimore: Johns Hopkins University Press.
- Wolff, Christian. 1732. *Psychologia empirica*. Frankfurt: Renger.
- Wolff, Christian. 1996. In *Discursus praeliminaris de philosophia in genere. Einleitende Adhandlung über Philosophie im allgemeinen; Historische-kritische Ausgabe*, ed. Gunther Gawlick and Lothar Kreimendahl. Stuttgart: Froomann-Holzboog.
- Zammito, John H. 2002. *Kant, Herder, and the birth of anthropology*. Chicago: University of Chicago Press.

Chapter 8

Back to the Facts – Herder on the Normative Role of Sensibility and Imagination

Anik Waldow

8.1 Introduction

In eighteenth-century Germany the battle cry “back to the facts” was usually heard when philosophers took issue with the speculative and experience-detached metaphysics standing in the tradition of Wolff and Leibniz.¹ Thus metaphysics was seen as an imaginative exercise that failed to account for observable facts, while system-building, deductive reasoning and a commitment to universal structures were identified as the hallmarks of useless philosophizing. Since metaphysics was here conceived as an intellectual activity that failed to use the mind correctly, namely, in such a way that the mind is able to constrain its speculative imagination, questions about the manner in which the mind can enter into an epistemic relation with the world became intimately related to questions about the norms of philosophy. In other words, what came under attack was not just one particular philosophical tradition, but the aims and methods of the entire discipline.

Interestingly, by the mid-eighteenth century literally no one wanted to be associated with the charges of experience-detached speculation. Even Wolff, whom many regarded as the champion of speculative metaphysics,² objected to the kind of imagination-driven methodology characteristic of Descartes’ fictitious vortex theory,³ thus aligning himself with the proponents of experiment and observation.⁴ Other

¹ See Zammito (2002, 15–41), Gaukroger (2010, 283–289) and Beiser (1987, 165–192).

² Schneider (1985, 58–92).

³ Wolff (1755, Sect. 3, 140–147). See Vanzo (manuscript) for a discussion of Wolff as a proponent of experimentalism.

⁴ Thus Wolff states in *Elementa matheseos universae*: “[2]We do not approve fabricated hypotheses, that many people nowadays ably introduce within natural philosophy, pretending anything arbitrarily, being altogether unable to prove what they imagine that could exist in nature”, Wolff (1741, vol. 5, § 492). Attacks on the use of hypotheses were usually mounted when expressing

A. Waldow (✉)

Department of Philosophy, University of Sydney, Sydney, NSW, Australia

e-mail: anik.waldow@sydney.edu.au

prominent thinkers followed suit. In his 1785 review of Herder's first volume of *Ideen zur Geschichte der Menschheit* (1784),⁵ Kant objects to Herder's so-called method of analogy,⁶ precisely for the reason that he deems this method speculative. Summarizing his dissatisfaction with Herder's lack of philosophical rigour, at the end of the review Kant urges that conceptual analysis, observable laws and systematic reasoning instead of unconstrained imagination amplified by metaphysical and emotion-laden speculation should guide Herder's work when completing the forthcoming volumes of the *Ideen*.⁷ Kant's contrast between imagination, emotions and sentiments, on the one hand, and reason and "cold judgement,"⁸ on the other, stresses the negative, philosophy-undermining effects of non-rational capacities, such as sensibility and imagination.

This essay will defend Herder against Kant in order to gesture towards a more positive conception of imagination and sensibility by arguing that both are compatible with rational inquiry, and in a specific sense even required for an anti-speculative engagement with the world. For Herder, this is the case because the imagination is a form of inner sensibility that focuses the mind on representations of particulars, that is, concrete instantiations of nature and human life. In this function, it prevents the mind from engaging in the kind of reasoning concerned with abstract universal structures and metaphysical speculation. The intention of developing this defence is not merely to show that Herder remains true to his own anti-metaphysical tenets. More generally, the aim is to improve our understanding of the nature and scope of experience-oriented inquiry. If it can be shown that a sensibility-based imagination is conducive to the generation of philosophical insights, it becomes clear that the kind of inquiry promoted in opposition to school metaphysics is not an observation-glorifying exercise. Rather, it is a position marked by the readiness to integrate non-rational capacities into a reason-guided examination of nature and human life, thus transforming our understanding of what reason is and what it ought to achieve.

I will first examine Herder's claim that all thinking derives from sentiment to show that on his account imagination plays a positive role in the formation of concepts and through this enables the mind to avoid meaningless speculation. The subsequent section will analyze Herder's polemics against metaphysics in relation to his demand that reason should work in unison with the mind's non-rational

one's alliance with non-speculative philosophy and stressing one's indebtedness to Newtonian experimentalism. An example of this kind is provided by Jean Le Rond D'Alembert in the *Preliminary Discourse to the Encyclopedia of Diderot*. He here objects that geometers remove their claims far "from what really exists in Nature" when they do not base their calculations on "appropriate experiments", D'Alembert (1995, 24).

⁵ Herder (1985), *Ideen zur Philosophie der Geschichte der Menschheit*, vol. 6, henceforth *Ideen*.

⁶ Kant writes, "Es soll mit Vermeidung aller metaphysischen Untersuchungen die geistige Natur der menschlichen Seele, ihre Beharrlichkeit und Fortschritte in der Vollkommenheit aus der Analogie mit den Naturbildungen der Materie vornehmlich in ihrer Organisation bewiesen werden," Kant (1902), vol. 8, 51 (henceforth AA).

⁷ AA, 8, 55.

⁸ AA, 8, 45.

capacities, thus revealing the tight connection between reflections on the norms of a commendable use of reason and the endeavour of turning philosophy into a world-focused discipline. The final section will discuss the projective character of the imagination in relation to the problem that the discovery of novelty requires the inquirer to reach beyond her familiar conceptual framework. The aim of this is to provide a concrete example of the manner in which a sentimental, imagination-driven engagement with the world is able to expand one's knowledge.

8.2 Concept Formation

In order to establish that for Herder reasoning is an operation of the mind that in fact requires a sensuous and imagination-driven engagement with the world, I will begin by analyzing Herder's account of the origin of language.

In the prize-winning essay *Abhandlung über den Ursprung der Sprache* (1772)⁹ Herder argues that human language emerges from the so-called state of *Besonnenheit* in which the mind is able to stand back from the incessant flood of sensations, thus becoming aware of its seeing, thinking and willing. Herder is clear that *Besonnenheit* is unique to humans whom he otherwise considers as hopelessly under-equipped creatures (he refers to humans as "Mängelwesen"): endowed with only a few rudimentary instincts and deprived of senses fitted to a specific environment,¹⁰ humans are particularly vulnerable and not very apt to secure their survival. Instincts are here analyzed as forces that mechanically focus a creature's attention on events happening in her environment, while the relative lack of instincts in humans is taken to enable them to concentrate on their own mental operations, thus creating an awareness of themselves as beings with sensations, desires and thoughts.¹¹

Herder argues that once in this conscious state of mind humans possess everything needed to render the emergence of language possible. Through the mind's awareness of its sensations it is able to isolate and name certain features by affixing sounds to them. Herder gives the following example in which the sound of a sheep is used to signify the bundle of properties constituting the perception of the sheep as a whole: "White, soft, woolly – his soul operating with awareness, seeks a characteristic mark – the sheep bleats!"¹² By conceiving of the process of naming in this way, Herder stresses that mere sensibility is not enough for the development of linguistic competence, because what ultimately puts humans in a position to pick out and name certain properties is a unique form of awareness, that is, the state of *Besonnenheit*.

⁹ Herder (1985), *Abhandlung über den Ursprung der Sprache*, vol. 1, henceforth *Ursprung*. F and the relevant page number of Herder (2002) will signal that I have used the English standard translation of a selection of Herder's writings instead of my own.

¹⁰ *Ursprung*, 770, F 128.

¹¹ *Ibid.*, 716–717, F 82–83.

¹² *Ibid.*, 723, F 88.

Herder thinks of the ability to reason in close relation to an agent's linguistic competence, stressing that there can be no thought without words: "If now it has been proved that not even the slightest action of his understanding could occur without a characteristic word, then *the first moment of taking-awareness [Besinnung]* was also *the moment for the inward emergence of language.*"¹³ Furthermore, Herder agrees with Kant that reason enables us to exert some sort of control because, according to him, reason emerges from *Besonnenheit*, that is, the very state that puts humans in a position to overlook the chaos of spontaneously caused sensations. Yet, for Herder reason is a skill that activates itself in its interaction with the sensuous and cannot be comprehended in isolation from sensibility and imagination, as I will explain below. Therefore reason cannot be used to produce the kind of "logical accuracy in the definition of concepts, or diligent distinction,"¹⁴ which Kant recommends in his review as an antidote to Herder's emotion-laden flights of the imagination, if this ordering of concepts entails that reason has to free itself from the influence of these faculties.¹⁵

Herder often refers to the imagination as a form of inner sensibility that unites the inputs of different sense modalities the moment we perceive them: "If in this way everything flows together out of our senses into the imagination, or however we want to call this sea of inner sensuality, and our thoughts, sensations, and drives swim and float upon it, has nature not woven anything further that unites them, that guides them? Certainly, and this is the *nerve structure.*"¹⁶ In its function to render the inputs of perception accessible, imagination (enabled by the nerve structure) thus emerges as a necessary ingredient in regular processes of representation and thought. And even stronger than that, since the ability to reason, together with one's linguistic competence, is taken to emerge through one's conscious awareness of one's sensations, reason cannot be conceived independently of that inner sensibility Herder attributes to the imagination. After all, according to Herder, the imagination is the locus where sensations are consciously perceived.

To accept that reasoning is grounded in pre-rational forms of representation is of course not to deny that reason, and more specifically the possession of language, transforms the manner in which humans interact with the world. The fox, Herder argues, might improve its survival chances by anticipating a course of action on the basis of his past experiences, that is, by associating Humean causes and effects in response to previous observations. However, at the same time, Herder points out that the fox will never be able to trump the cunning of the hunter. The reason for this

¹³ *Ibid.*, 770, F 128.

¹⁴ AA, 8, 45, my own translation.

¹⁵ While Kant castigates Herder's uncontrolled use of the imagination in the review, he acknowledges the synthesizing force of the imagination as a valuable ingredient in thought formation in other places. For an "imagination-friendly" interpretation of Kant see Longenesse (1998); for an account that draws a sharp line between Kant's imagination-involving aesthetic judgment and imagination-free scientific judgment see Wenzel (2001).

¹⁶ Herder (1985), *Vom Erkennen und Empfinden der menschlichen Seele*, vol. 4, henceforth *Erkennen*, 350, F 205.

is that for linguistic creatures the “succession of ideas”¹⁷ takes on a new form. Presumably, this means that humans are able to combine and analyze their ideas in such a way that choices become available that take them beyond the mere reproduction of previously experienced patterns of action.

This positive account of reason as the very capacity that opens up new avenues of successful action might be taken to suggest that *after* the formation of language Herder thinks of reason as operating independently of the inner sensibility related to the imagination. Charles Taylor is a proponent of this view and writes:

Animals respond to natural and accidental signs (e.g. smoke is an accidental sign of fire, and clouds of rain). Humans have also “instituted” signs. The difference lies in the fact that by means of these latter humans can control the flow of their own imagination, whereas animals precisely follow those which are triggered off in them by the chain of events. There is obviously some link between Herder’s description of our interrupting the “ocean of sensations” and this Condillaquian idea of taking control.¹⁸

Against this reading, the objection must be raised that Herder never suggests that with the possession of language it becomes possible to detach oneself from the intermingling chaos of the imagination: naming does not proceed via the allocation of one particular sound to one specific sensation, thus enabling reason to establish its own clear-cut order. Rather, names come into existence in a situation in which sensations are undistinguished: “The soul, which stood in the throng of such a confluence of sensations, and in need of forming a word, reached out and got hold perhaps of the word of a neighboring sense whose feelings flowed together with this one. In this way words arose from all the senses.”¹⁹ Hence, it is an illusion to believe that words, and ultimately language and reason, enable the mind to overcome the chaos of intermingling sensations represented in the imagination. After all, the words and concepts through which reason operates rest on this intermingling chaos, thus suggesting that, instead of working against it, the order of reason supervenes on the order of the imagination.²⁰

This point becomes even clearer when looking at Herder’s account of abstraction. Abstract concepts, Herder concedes, are more orderly than those that bear a direct link with the realm of the sensuous in that they isolate certain features at a primary sensory level, as Sonia Sikka has pointed out.²¹ However, Herder also argues that this ordering of primary inputs of the senses possesses a tendency to divest concepts of “the multiple and sensuous”²² such that abstract concepts turn into “empty nutshells,”²³ “wooden walls and colorful houses of cards”²⁴ suitable for

¹⁷ Ursprung, 773, F 130.

¹⁸ Taylor (1991, 50).

¹⁹ Ursprung, 745, F107.

²⁰ Heinz captures this point when speaking of the mind’s obligation to avoid the suppression and death of originally lively sensory inputs and sentiments.

²¹ Sikka (2007, 39).

²² Herder (1985, *Metakritik*, vol. 8, 422).

²³ *Ideen*, 349.

²⁴ *Ibid.*, 341.

nothing but metaphysical play. In recognizing this threat of abstraction, Herder follows Locke who famously claims in response to scholastic procedures of formalized reasoning that philosophical discourse turns into meaningless dispute when based on concepts that lack a connection with experience-derived ideas.²⁵ In the *Epistle* of the *Essay* Locke writes:

Vague and insignificant Forms of Speech, and Abuse of Language, have so long passed for Mysteries of Science; And hard or misapply'd Words, with little or no meaning, have, by prescription, such a Right to be mistaken for deep Learning, and height of Speculation, that it will not be easie to persuade, either those who speak, or those who hear them, that they are but the Covers of Ignorance, and hindrance of true Knowledge.²⁶

It thus emerges that for Herder, as much as for Locke, the ordering of one's thoughts is desirable only if such ordering does not involve the kind of abstraction that removes concepts from their sensuous grounding. Since in Herder's case sensations become consciously available only once they have been synthesized by the imagination,²⁷ this entails that the aim of orderly, reason-guided thinking cannot consist in suppressing the influence of the imagination and the way it presents our various sensory inputs as united, because it is precisely this unity of one's sensations that provides the original grounding urgently needed to guard against meaningless speculation.

In order to understand what Herder's reason does when it engages in the generation of knowledge it is useful to focus on the difference between analysis and synthesis as two methods of handling sensory inputs. For Herder "Erkenntnis" is the product of the attempt to unify (*Vereinheitlichung*) the multiple at various levels of representation, as Marion Heinz has put it.²⁸ It is therefore not analysis, that is, the separation of constituent ideas that the imagination represents as united, but synthesis in the sense of finding a characteristic feature under which the multiple can be subsumed that enables knowledge.²⁹ In the light of this, we can see that the imagination's way of presenting sensory inputs as united should not be conceptualised as a flaw that must be corrected; rather this function of the imagination should be conceived as providing the mind with the opportunity to use its reason in order to find even greater unity. But as the starting point for reason's attempt to find even greater unity.

Although Herder conceives of reason as an original capacity in the form of a special reflective awareness (*Besonnenheit*), he nevertheless stresses that a state of fully-fledged reflection (*Besinnung*) is the product of one's sensuous engagement with the world.³⁰ This in principle already follows from Herder's claim that there

²⁵ For a comparison of the pragmatic intentions of Locke's and Herder's critiques of metaphysics see Sikka (2007, 35).

²⁶ Locke (1975, 10). Locke also identifies prejudice as a result of the "Abstractness of the Ideas" (*ibid.*, 8) figuring in one's thoughts, thus building a bridge between the Baconian notion of speculation as the result of an unreflective acceptance of delivered opinions and the abuse of language conceived as the juggling of empty meaningless formulas.

²⁷ See Heinz (1994, 159).

²⁸ *Ibid.*

²⁹ *Ibid.* and 130–131. Also see Heinz (1994, 166–170), for a discussion of the way in which analysis can be useful as a means of driving out differences already present in sensory experiences.

³⁰ Ursprung 770, F 128.

can be no thought without a word. After all, if it is the case that thought is impossible without language and, moreover, that language evolves in relation to one's interactions with concrete environmental patterns (for instance by encountering sheep and naming them), then it would not be possible for us to reason before we enter into a sensuous relation with the world, because it is this relation that gives rise to the concepts necessary for reason to get off the ground in the first place.³¹ Furthermore, since on this account language emerges in relation to interactions with one's surroundings, that is, an environment that in some cases might contain sheep while in others it might not, Herder's conception of reason recognizes the variability of different manners of thought in relation to differences in those environmental patterns constitutive of the formation of an individual's ability to speak. He thus offers a naturalistic explanation of socio-cultural diversity in language and reason.³²

I will return to an examination of the relation between nature and culture in Herder in Sect. 8.4 of this paper when discussing how an imaginative engagement with culturally variable forms of conceptualization can be conceived as enhancing a better understanding of the world. Before entering into this discussion, however, I want to return to my critique of Taylor's claim that Herder adopts the Condillan idea of taking control. For Condillac it is in virtue of language that the mind is able to focus its attention on the things that matter in its sensation.³³ Furthermore, by claiming that persons can "make signs for themselves only when they live together,"³⁴ he stresses that commerce with other persons is a necessary condition for the invention of linguistic signs. Squared with one another, these two claims suggest that perfection of thought (in the sense of being more focused on what matters to one's sensations) requires an artificially induced reform of language, because it is language and the application of conventionally established signs that enable the mind to control and order its sensory inputs.³⁵

As we have seen, Herder takes a rather different course. He accounts for the origins of language by focusing on the situation of *one* individual in a specific environment. He cites encounters with sheep and their sounds as the starting point for the ability to name and conceive of these animals. So, unlike Condillac, Herder thinks that social collaboration is not the root of our linguistic competence but

³¹ Nigel DeSouza stresses that for Herder reason is not "a further capacity that is *added* to the capacities we share with animals and that requires other 'arbitrary or social forces' to develop," DeSouza (2012a, 223). DeSouza is here arguing that for Herder reason and language *necessarily* develop given that humans are put in the state of "Besonnenheit" that no other animal experiences. He is not denying that on Herder's account the activation of reason and language requires humans to interact with their environment. In this context he speaks of the "flooding of the senses by the external world" as the cause that "activates the individual's sensorium commune," *ibid.*, 227.

³² Herder develops his theory of culturally diverse manners of thinking in *Von der Veränderung des Geschmacks*, Herder (1985), vol. 1; henceforth *Geschmack*.

³³ Condillac (1746, part I, § 2, Chap. 4, 46).

³⁴ *Ibid.*, 65.

³⁵ The importance of language in enhancing the mind's ability to think properly is a theme that surfaces in many of Condillac's earlier and later writings; see Condillac (1746, part II, § 1, Chap. 15, 200–201), Condillac (1780, 109–110, 137, 168), and Condillac (1991, Chap. 14, 247). See Riskin (2002, 240), for a discussion of this point.

instead an uncompromised engagement with the world as it presents itself to us in an unordered mix of sensations.³⁶ If we now recall that his critique of abstraction suggests that a link to one's sensibilities ought to be maintained in order to prevent the degeneration of words into meaningless concepts, it becomes clear that for Herder an artificially reformed language that removes and reorganizes links with the sensuous can never have the normative standing it has for Condillac. If such an artificial language took control of our thoughts, and in a Condillaquian fashion "helped" us to organize the chaos of intermingling sensations and sentiments, it would remove us from the very grounding that renders our concepts and reason meaningful.

8.3 Herder's Holism

So far I have argued that Herder regards sensibility and imagination as the natural ingredients of reason and thought – ingredients that he deems crucial if one's concepts are to remain meaningful at all. The aim of this section consists in providing further support for the claim that the imagination ought to be part of cognitive processes. While up to this point we have mainly focused on its role in ordinary processes of representation and concept formation, the subsequent discussion will refer to the imagination as a capacity that is able to take us beyond the limits of sense perception. In order to reveal that even in this role imagining can valuably contribute to an experience-focused engagement with the world, I will first consider the common claim that experience-detached reasoning is in itself a form of imagining, before suggesting that it makes sense to distinguish two kinds of imagination: one that is analogous to sense perception and conducive to an experience-informed engagement with the world; and another that is associated with the sort of exaggerated, cold reasoning common in metaphysical speculation. By focusing on this difference, it will be revealed that speculation is the product of the mind's failure to control reason by integrating sensibility and imagination in its thinking.

Whenever Herder attacks school metaphysics, he objects to useless methods of teaching, empty definitions and formal logic for their counterproductive effect on free thinking. Echoing Rousseau, Herder writes: "Why do we have so few independent thinkers? Because already in school they were hemmed in with Logic."³⁷ Upholding the German ideal of *Bildung*, Herder requires teachers to encourage the student's active engagement with the subject at hand: "Set before him instead of

³⁶ Heinz remarks in this context that for Herder there is no contradiction involved in thinking of language as occasioned by nature and as man-made at the same time, because it is not the way the world objectively is but the way it is experienced that enables humans to name it; see Heinz (1994, 163).

³⁷ Herder (1985), *Wie die Philosophie zum Besten des Volks allgemeiner und nützlicher werden kann*, vol. 1, 111, F 9; henceforth *Wie die Philosophie*.

words a large number of actions, let him see instead of reading, instead of wishing to educate [*bilden*] his head let him educate [*bilden*] himself.”³⁸ The exercise of passively memorizing, Herder continues, should be abandoned to make room for a student’s active resonating with the world. Learning here takes the form of an original reception that requires openness and responsiveness as much as an unprejudiced readiness to experience the world.

These polemics bear striking similarities to Bacon’s critique of school metaphysics in *The Advancement of Learning* (1605). Herder and Bacon both critique the *status quo* of learning and philosophy in general as a hopelessly speculative enterprise in view of its disengagement from the realm of experienceable particulars. Furthermore, both regard the benefits of an active engagement with the world as the first step in a process that enables inquirers to free themselves from dogmatic inculcation. Bacon’s discussion of three of the four human idols clearly illustrates his concern with undue influences on the mind’s thinking. Thus, he presents the *idols of the cave* as resulting from an individual’s “upbringing” and “company” as well as from “the reading of books” and the uncritical acceptance of “authority.”³⁹ The *idols of the marketplace* are described as illusions following from the mind’s readiness to take over from others “a poor and unskillful code that obstructs the understanding,”⁴⁰ while the *idols of the theatre* are identified as “illusions which have made their homes in men’s minds from the various dogmas of differences of philosophies, and even from mistaken rules of demonstration.”⁴¹ The general message of this discussion is that speculation is the product of a passive and absorbing state of mind that needs to be replaced by an active curiosity-driven engagement with the world.⁴² Hence, for Bacon as much as for Herder the relevant contrast is not between reason and imagination, but between self-responsible experience-focused inquiry and other-dominated indoctrination.

To bear this contrast in mind is important, because it complicates the widespread idea that speculation emerges when the mind strays away from evidence provided by observation. If speculation arises as the result of an uncritical acceptance of philosophical dogmas and inculcation that dovetails with an exaggerated commitment to experience-detached forms of abstract formal reasoning,⁴³ it is possible to count claims about matters of fact as non-speculative despite the impossibility of directly observing what is claimed to exist. The reason for this positive evaluation of claims about the unperceived is that the criterion defining speculation is not the

³⁸ *Ibid.*, 126, F 22.

³⁹ Bacon (2002, 41).

⁴⁰ *Ibid.*

⁴¹ *Ibid.*, 42.

⁴² Note that for Bacon a successful handling of the idols requires the support of instruments, especially because the first class of the idols, namely the *idols of the tribe*, are natural to all human beings; see Bacon (2002, 41). See Waldow (2010) for an account that discusses Bacon’s motives for invoking the specter of skepticism.

⁴³ See Gaukroger (2001, 6–67), for an account that comes at the same point by focusing on the function of the concept of the persona of the inquirer.

failure to observe the things in question, but the involvement of abstract formal reasoning detached from *any* kind of observational evidence.

A closer look at one of Kant's major objections to Herder may be helpful to elucidate what is at stake. Kant is particularly critical of the concept of reason that Herder develops in the *Ideen*. Herder here thinks of reason as a special instantiation of the organic force that pervades the entire universe and is responsible for change and development.⁴⁴ Kant objects that this conception not only misrepresents what reason is, namely something that cannot be found in nature, but also signifies a return to the belief in occult metaphysical powers that can nowhere be perceived.⁴⁵ He asks:

What shall we think of the hypothesis of invisible forces that bring about the organisation of different stages [of life], a hypothesis of the kind aiming to explain the unknown by making reference to the even more obscure? While we can study the laws of the one through experience, however without penetrating into their causes, we lack experiences with respect to the other. And what can be cited in support of the philosopher's approach apart from the sheer despair to gain insight from some unfounded claim about nature and the deliberate decision to make use of his creative imagination?⁴⁶

⁴⁴ These forces must be conceived as souls at the level of the organic world, and as responsible for the diversification of life; see DeSouza (2012b, especially 786–788), for a helpful discussion of Herder's concept of reason as a causally efficient version of Leibnizian powers. This conception of the soul as a force that interacts with the body for the purpose of maximizing the benefits of one's embodied existence is already present in his prize-winning essay about the origins of language. Herder writes, "Man nenne die ganze Disposition seiner Kräfte wie man wolle, Verstand, Vernunft, Besinnung u.s.w. Wenn man diese Namen nicht für abgesonderte Kräfte, oder für bloße Stufenerhöhungen der Tierkräfte annimmt: so gilt mir gleich. *Es ist die 'ganze Einrichtung aller menschlichen Kräfte; die ganze Haushaltung seiner sinnlichen und erkennenden und wollenden Natur;*" oder vielmehr – Es ist die *'einzige positive Kraft des Denkens,* die mit einer gewissen *Organisation des Körpers* verbunden bei den Menschen so *Vernunft* heißt, wie sie bei den Tieren *Kunsthähigkeit* wird." Ursprung 717, F 83. At the same time, however, Herder explains that reason is something that cannot be conceived as coming into existence by developing one's animal instincts further, but instead must be taken to exist in us as an active force (Besonnenheit) right from the start of our lives. This conception of reason, as DeSouza has argued, goes against Rousseau who would take reason to evolve out of instincts; see DeSouza (2012a and note 27).

⁴⁵ The relevant passage reads: "Was nun aber jenes unsichtbare Reich wirksamer und selbstständiger Kräfte anlangt, so ist nicht wohl abzusehen, warum der Verfasser, nachdem er geglaubt hat aus den organischen Erzeugungen auf dessen Existenz sicher schließen zu können, nicht lieber das denkende Princip im Menschen dahin unmittelbar, als bloß geistige Natur, übergehen ließ, ohne solches durch das Bauwerk der Organisation aus dem Chaos herauszuheben; es müßte denn sein, daß er diese geistigen Kräfte für ganz etwas anders als die menschliche Seele hielt und diese nicht als besondere Substanz, sondern bloß als Effect einer auf Materie einwirkenden und sie belebenden unsichtbaren allgemeinen Natur ansähe, welche Meinung wir doch ihm beizulegen billig Bedenken tragen," AA, 8, 53.

⁴⁶ AA, 8, 53. Here is the original: "Allein was soll man überhaupt von der Hypothese unsichtbarer, die Organisation bewirkender Kräfte, mithin von dem Anschlage, das, was man nicht begreift, aus demjenigen erklären zu wollen, was man noch weniger begreift, denken? Von jenem können wir doch wenigstens die Gesetze durch Erfahrung kennen lernen, obgleich freilich die Ursachen derselben unbekannt bleiben; von diesem ist uns sogar alle Erfahrung benommen, und was kann der Philosoph nun hier zur Rechtfertigung seines Vorgebens anführen, als die bloße Verzweifelung den Aufschluß in irgend einer Kenntniß der Natur zu finden und den abgedrungenen Entschluß sie im

Pointing to the absence of experiences that could support the belief in the existence of Herder's life forces, while accusing Herder of falling prey to metaphysical speculation, Kant gives us the impression that Herder's account is crucially lacking in observation and experimentation and for this particular reason qualifies as an instance of metaphysical speculation and dogmatism.⁴⁷

Yet, Herder's reference to active forces in nature does not stand without the support of the experimental sciences of his time.⁴⁸ Buffon's theory of reproduction and Haller's theory of irritability both challenged the concept of dead matter.⁴⁹ Herder was familiar with their writings and often developed his own theories by aligning with some of their central claims.⁵⁰ For instance, in *Vom Erkennen und Empfinden der menschlichen Seele* (1778) Herder conceives of the body as a machine that is never at rest, pulsating and vibrating in its incessant involvement with the world,⁵¹ adapting insights gained from Haller's experimentation with muscle fibers. Thus, perceptual processes are regarded as an expansion and contraction of sensibility brought about by the contact with certain perceptual stimuli,⁵² while the sophistication of bodily systems is understood as resulting from complex irritabilities contained in the fibers of their subsystems.

fruchtbaren Felde der Dichtungskraft zu suchen? Auch ist dieses immer Metaphysik, ja sogar sehr dogmatische, so sehr sie auch unser Schriftsteller, weil es die Mode so will, von sich ablehnt." Kant makes a similar point about the imperceptibility of soul forces in his letter to Mendelsohn, 8 April 1766, quoted in Reich (2001 355–356). See Heinz (1994, 28–30) for a discussion of Kant's critique in relation to his essay "Träume eines Geistessehers."

⁴⁷ In the preface of the *Ideen* Herder straightforwardly rejects the existence of occult metaphysical powers; see *Ideen*, 17.

⁴⁸ Sikka has claimed that Herder's pantheism could in principle be regarded as a form of Spinozistic naturalism; see Sikka (2007, 49, note 70).

⁴⁹ Buffon (1750) and Haller (1757–1766, 1981).

⁵⁰ Although Herder was familiar with their writings, it is difficult to determine what kind of naturalist Herder was. Lewis Spitz for instance argues that he was more of a creationist than a transformist; see Spitz (1955, 464). He bases his analysis on Henry Nevinson's *A sketch of Herder and his Times* (1884, 357), a work that would reveal Herder as standing in between Spinoza and Darwin. Eugen Sauter offers an analysis of the differences between Herder and Buffon; see Sauter (1910, 91).

⁵¹ See *Erkennen*, 346–347 and 201–202. Herder's conception of the human body as the locus of sensibility enables him to assert an uncompromised realism, see Sikka (2007, 42). According to him, the human body is embedded in the world, acts and interacts with it, thus furnishing the mind with information necessarily required for any form of understanding. To illustrate this point Herder speaks of the so-called "Saitenspiel" in *Erkennen*, 351, F 206. By treating sense perception as the effect of our pulsating with the universe, Herder is clear that we have no other choice than to trust our senses (*Erkennen* 348, F 203), precisely because all reasoning is made possible through the body's affectedness. See Zeuch (1998) for a critique of Herder's unproven assumption "that the world orients itself in accordance with our subjective representations," 152.

⁵² Heinz thinks of the dependence of the soul on the body as one of the fundamental tenets of Herder's concept of mind, "Grundlage der herderschen Argumentation ist der Gedanke, dass eingeschränkte Wesen die Inhalte ihres Vorstellens nicht durch oder aus sich selbst haben, sondern vermittels durch den Körper. Die Kraft der Seele ist daher gebunden an die Kraft des Körpers, ihre Wirkung ist essenziell ein Wirken auf die Körperkräfte und ein Bearbeiten dessen, was ihr der Körper zuströmt," Heinz (1994, 139).

We can here see that not everyone unanimously agreed on the kinds of empirical evidence required to guard against metaphysical speculation. Although opponents of metaphysics more or less converged on the opinion that in one way or another observation was necessary to avert the charge of speculative metaphysics, it was not clear how much and what kind of observation sufficed for this purpose. The flipside of this uncertainty concerning the role of observation was that it was also far from clear at what point exactly reasoning turned into the kind of formalized, rule-guided juggling with meaningless concepts that marked out theories as hopelessly metaphysical.

Herder conveniently circumnavigates these questions by arguing for a holistic concept of mind. In *Erkennen und Empfinden* he repeats his earlier claim that memorizing ought not to be the aim of learning, objecting to the “verbal memory of the school pedants”⁵³ for its tendency to turn the soul into a dry register of names, just before asking: “But for a *Caesar* or *Mithridates* was not in these cases their *memory for names* essential too?”⁵⁴ By juxtaposing the negative and positive aspects of memorizing in this way, Herder demonstrates the need for an integrated use of the mind’s various capacities: “All of these forces are at bottom only a single force if they should be human, good, and useful – and this is *understanding, intuition* with inner *consciousness*. Let one remove this from them, and the imagination is illusion, the wit childish, the memory empty, the cleverness a cobweb.”⁵⁵ These powers, Herder argues, become enemies if left separate, but enhance and promote each other when united to one single force: “Memory and imagination become the extended and deep image of truth; cleverness separates and wit combines so that precisely a clear weighty *One* arises.”⁵⁶

Thus, for Herder neither the imagination nor reason nor the senses possess value in themselves. Imagination detached from the intellect and unchecked by our sensibilities is as bad as reasoning that only obeys its own rules, because both enhance speculation and illusion. With this focus on how the mind ought to proceed, Herder’s descriptive analysis of how the mind works in principle transforms into a normative account. He argues that the mind’s natural propensity to engage in an integrative use of its capacities is valuable and must be sustained, thus turning the mere fact that “sight and hearing decode each other reciprocally”⁵⁷ into a cognitive virtue that helps the mind to achieve its best results.

By following Herder’s reflections on the need to synthesize the mind’s various powers in combination with his polemics against speculative metaphysics, we can now see that the burning question within the debate about the misguidedness of philosophy is not so much that of how sense perception can be conceived as an avenue to knowledge that competes with reason-dominated approaches as it is a question of how a reason that is lifeless and unable to engage with the reality of

⁵³ *Erkennen* 357, F 210–211.

⁵⁴ *Ibid.*

⁵⁵ *Ibid.*

⁵⁶ *Ibid.*

⁵⁷ *Ibid.*, *Erkennen* 349, F 204.

human experience can be cured.⁵⁸ In the previous section I argued that on Herder's account such a cure is possible if the mind operates with concepts that remain linked to the realm of the sensuous. The ability to use reason properly was thus evaluated in relation to the concepts figuring in processes of reasoning. In this section I have approached questions about the norms of thought by reflecting on the mind's different capacities. This discussion suggested that for Herder the manner in which the mind ought to proceed requires the mind to integrate its various capacities, because only then can it prevent its reasoning from running wild and tuning into a species of the imagination that fails to account for the way the world affects us. This result once more stresses that for Herder reason cannot be regarded as an isolated faculty that in some cases operates with the wrong kind of concepts, but must rather be conceived as a mental activity that unites a great mix of rational *and* non-rational capacities.

Here, then, it becomes clear that the role of the imagination within debates about the aims and goals of philosophy is not that of a seducer leading reason away from its proper task. As we have seen, when proponents of experience-focused forms of philosophy antagonize against school metaphysics, they usually refer to the imagination as a form of reason that fails to generate useful and empirically grounded explanations. This failure occurs when sensations and the way they are represented by our inner sensibility – which Herder also calls the imagination – do not feed into our reasoned thoughts. So rather than seeing imagination as the cause of illusion, it is treated as the *product* of failures to integrate one's sensibility into processes of philosophical thought. Herder, as has been discussed, endorses the view that reason should act in unison with sensibility and imagination. By allowing our non-rational capacities to play this role, he in principle legitimizes contributions of the imagination that go beyond its mere synthesizing effects in representation as long as these contributions are well integrated with other cognitive processes. In the next section I will specify the particular manner in which imagination can help us to generate knowledge by focusing on its ability to enact experiences we cannot have because of our contingent location in space and time.

8.4 Imagining as a Form of Discovery

While so far we have mainly looked at the imagination as an ingredient in relation to processes of reasoning, I will now specify the conditions under which imagining can usefully supplement sense perception and, through this, enable an expansion of knowledge. For this purpose I will examine Herder's use of the German word

⁵⁸ It is interesting to note in this context that in the eighteenth century life was associated with synthesis. Jessica Riskin writes, "Synthesis was the process of life; analysis could capture only death," Riskin (2002, 247). Against this background, Herder's push for an integrated use of the mind's powers can be regarded as yet another attempt to conceive of human reason as a natural life force; see note 39.

“Einfühlung,” which denotes the ability actively to feel one’s way into the reality of a certain person or situation.

In *Von der Veränderung des Geschmacks* (1766) Herder emphasizes that conventions shape our language and reason. He here treats thinking as a culturally conditioned manner of thought (Denkart), comparing it to taste, which varies in accordance with an individual’s time and place. At first sight, it may look as if this conventionalist account of language gives rise to conflict. As has been pointed out, in *Abhandlung über den Ursprung der Sprache* (1772) Herder develops a naturalistic account of the origin of language: he explains that an individual’s ability to form thoughts derives from her interactions with her natural environment.

To see that both accounts are perfectly compatible, however, we need only remember that in the *Abhandlung* Herder is concerned with the question of how it is possible for pre-linguistic creatures to develop language, while in the earlier essay on taste he refers to a situation in which language has already emerged. The absence of linguistic conventions in his treatise on the origin of language thus explains why he ignores what he has urged earlier and repeats in the much later *Ideen* (1784), namely, that encounters with linguistically organized social environments and their conventions are in fact constitutive of an individual’s ability to think and reason.

As explained in Sect. 8.1 when discussing Herder’s account of the *Abhandlung*, his genealogy of linguistic conventions suggests that they are the product of interactions between a specific group of people with environmental patterns characteristic of their natural and social situation. If we now take into account that, once language has emerged, these interactions involve confrontations with language, and *per extension* specific manners of thought, it appears that cultural differences engendered by one’s being trained within a specific linguistic community bear a link to both an individual’s cultural *and* natural situatedness. This double link is explained by the fact that the conventions themselves reflect the environmental conditions of those individuals who were involved in shaping them. The relevance of this claim will become clear presently when I explain why an engagement with foreign manners of thinking does not merely diversify one’s conceptual resources but also enhances one’s grasp of reality. Before turning to this issue, however, it is important to uncover the relevance of non-rational capacities in rendering accessible the obscure and unknown.

Given the situatedness of one’s thinking as a result of one’s location in a specific historical and cultural context, Herder argues that an engagement with foreign manners of thought is a difficult task that all too often leads to projection rather than the detection of the genuinely new.⁵⁹ To illustrate this point, Herder refers to Shaftesbury’s failure to engage with Homeric society. Because of Shaftesbury’s propensity to project the values of his own time onto Homer’s writings, he fails to understand the pedagogical use of Homeric poetry in antiquity.⁶⁰

Although objecting to Shaftesbury’s hermeneutical shortcomings, Herder recognizes the difficulty the mind experiences in seeking to overcome its accustomed

⁵⁹ *Geschmack* 157–158, F 253–254.

⁶⁰ Herder (1985), *Über die neuere deutsche Literatur. Zwote Sammlung von Fragmenten 1767*, vol. 1, 321–322, F 46–47; henceforth *Fragmente*, zweite Sammlung.

ways of conceptualizing events as something natural. Arguing, on the one hand, that all language is culturally determined, while, on the other, that language determines our manner of thought, Herder arrives at the following question: if we are culturally determined to think the way we do – which is precisely the reason why we tend to project rather than detect – how is it possible to enter into new manners of thinking and to learn something previously unknown? Or, more concretely put, the question is: how is it possible to learn, for example, the true message of Homer if we naturally tend to block cognitive access to the unknown and novel by projecting our familiar patterns of thought onto everything new?

Herder's account of mythology offers an answer to this question. By stimulating the imagination, he argues, mythology renders accessible that from which our training in a specific manner of thought usually bars us. In this context Herder speaks of "die Bildung unserer Erfindungskraft,"⁶¹ thus emphasizing the importance of utilizing this capacity for the purpose of learning. Herder writes:

Here we listen to the Greeks, how their poetic imagination, their sensuous thinking was artful enough to dress the truth in images ... Now that we are surrounded by a new world of discoveries: poets among us, let us taste the mighty honey of the ancients ... Learn from them the art of creating images for each one of us in our own private and unique sphere. Instead of paling in the face of that appalling image that Homer spit out, firm up your mind to drink from the ocean of images and particularities that surrounds you.⁶²

The thought which legitimizes this non-rational approach towards learning is that since the mind tends to obstruct its learning by simply projecting its own familiar conceptual structures onto newly encountered contexts, an approach is needed that would connect the mind with the new while circumventing the involvement of too many conceptual resources. The passage above suggests that such an approach must consist in an engagement with novelty that stimulates imagination and sentiment, that is, an engagement that allows the mind to immerse itself in new experiences.⁶³ To emphasize this aspect of direct acquaintance, Herder invokes the metaphor of

⁶¹ Fragmente, zweite Sammlung, 447. Interestingly, the German word "Erfinden" is as close to the English "imagining" as it is to the word "invention," thus stressing the valuable discovery-conducive aspects of an imaginary engagement with novelty.

⁶² "Hier belausche man die Griechen, wie ihre dichterische Einbildung zu schaffen, wie ihre sinnliche Denkart, abstrakte Wahrheit in Bilder zu hüllen wusste ... Und da wir in eine neue Welt von *Entdeckungen* um uns haben: ihr Dichter unter uns, so kostet von dem mächtige Honig der Alten ... Lernet von ihnen die Kunst, euch in eurer ganz verschiedenen Sphäre eben so einen Schatz von Bildern verdienen zu können. Statt, dass ihr, nach jenem ekelhaften Gemälde, das, was Homer gespien hat, euch beliebt lasset: so stärkt euer Haupt um aus dem Ozean von Erfindungen und Besonderheiten, der euch umfließt, zu trinken," Fragmente, erste Sammlung, zweite überarbeitete Ausgabe 1768, 449.

⁶³ To be precise, the passage continues as follows: "Ich meine statt dass ihr aus den Alten Allegorien klaube, oft wo sie gewiss nicht daran gedacht; so lernet von ihnen die Kunst zu allegorisieren," *ibid.* This might be read as suggesting that the intention of immersing oneself in Homeric poetry is to learn a certain technique rather than to understand a specific aspect of Homeric society. Learning this specific technique, however, does in fact require one's entering into the images and sentiments that Homer's writing elicit. So it still holds that for Herder learning that enables the mind to surpass its accustomed ways of thinking (and writing) requires the activation of non-rational capacities.

drinking from the ocean of images and particularities, thus reaffirming the idea that learning should never take the form of a lifeless repetition of formulas but ought to involve a lively resonating with the world.

It stands to reason that the ability to experience images and sensations still involves some kind of conceptual resource. After all, Herder himself defines the imagination as the locus where words and signs flow together with images, sounds, and emotions.⁶⁴ However, according to what was said in Sect. 8.1, where the role of sensuousness was discussed in relation to the meaningfulness of concepts, it would appear that his invocation of mythology as an avenue to new manners of thinking is not intended to rule out the involvement of *any* conceptual resources but rather merely to highlight the limitations of an overly theoretical and abstract approach towards the newly encountered.⁶⁵

Furthermore, as we have seen, Herder conceives of the sensuous as the ground from which concepts arise, which is why he is so confident that sentiments and images can in fact facilitate learning. If sensuous encounters enable the formation of thought, it can be expected that new combinations of images, feelings and sensations triggered by an imaginative engagement with the unknown will likewise yield new ways of thinking about the world.⁶⁶ Thus, within processes of thought, images and emotions function as important building blocks through which we can enrich our conceptual resources, broaden our perspective and generate a better grasp of reality.

It is crucial to realize at this point that more is at stake than the question of how to multiply one's way of thinking about one and the same thing. As has been pointed

⁶⁴ *Erkennen* 349, F 204.

⁶⁵ A more reason-guided way of entering into new manners of thinking surfaces in Herder's account of translation. Each language, Herder claims, is organized upon its own principles, possesses a unique grammar and syntax, has its own metaphors and inversions and expresses meaning in very different ways; see *Fragmente, zweite Sammlung*, 199–205, F 37–42. Given this diversity, translators are bound to fail when trying to press one language into another. They should therefore restrict their efforts to incorporating some of the foreign structural elements into their own language to diversify the corpus of their expressions and thoughts. Languages that are most suitable for this task are those resembling one's own; see *Fragmente, zweite Sammlung*, 205; F 43. This last point about resemblances suggests that even when actively trying to penetrate into novelty it is one's intuitive grasp, made possible through certain resemblances with the familiar, that supports most effectively one's learning.

⁶⁶ Herder writes that “der Europäer hat keinen Begriff von den heißen Leidenschaften und Phantomen, die in der Brust des Negers Glühen, und der Inder keinen Begriff von den unruhigen Begierden, die den Europäer vom einen Weltende zum anderen jagen,” *Ideen*, 331. One could interpret this as a statement about the impossibility of accessing the mental dimension of other persons in order to stress that Herder is more sceptical about the benefits of the imagination than has been suggested. To refute this reading, we only need to consider that in the cited passage Herder refers to the absence of *concepts*, namely those required for an understanding of the other person's sensibilities. By raising the issue in this way, he in principle confirms what has been said before. That is, that an engagement via emotions and images is necessary to equip the mind with the concepts currently missing from its culturally determined conceptual framework, but which are required to enable a comprehensive understanding of certain aspects of human mental life. As such, nothing in this passage challenges the claim that an imaginative engagement with novelty is able to create new knowledge.

out, for Herder diverse manners of thinking manifest themselves in diverse linguistic practices, and must be seen as coming into existence through a group's interaction with a specific environmental pattern. Given this connectedness between the way one conceptualizes the world and the way one interacts with it, it follows that an engagement with a hitherto unknown way of conceptualizing reality can give access to a hitherto unknown way of being immersed in a specific environmental milieu. A change in perspective brought about by one's engagement with a different conceptual framework via one's sentimental responses can thus be taken to lead to the discovery of something new, namely, those aspects of reality that are unknown to us but are constitutive of the experience of people living in conditions different from our own. A sentimental and imaginative engagement with foreign manners of thinking is therefore not only desirable for the purpose of diversifying one's conceptual resources in relation to things that we have encountered already, but also in relation to the discovery of novelty.⁶⁷

Conceived in this way, it becomes clear that in Herder's conceptual framework the imagination performs a similar role to sense perception. After all, as has been argued in the previous section, it is the imagination that grounds thoughts via our inner sensibility in concrete representation of human life and nature, thus blocking the mind's tendency to get lost in empty abstract formulas that only marginally relate to the way in which we experience reality. In this function, the imagination effectively cancels out the kind of speculation typically associated with school metaphysics. Furthermore, as has been argued in this section, the mind also relies on its imaginative capacities when it projects itself into different times and places and so renders accessible experiences it cannot acquire through the senses. By multiplying one's experiences in this way, the imagination thus compensates us for the limited sense perception to which we are confined by the contingency of human existence.

To be sure, this compensation only succeeds if the imagination works within well-defined limits, allowing for an integration of what is imaginatively encountered into one's existent framework of beliefs, while also being supported by direct experience and reason. Herder's holistic concept of mind expresses this demand rather clearly, as does his account of translation. Translation, he argues, is a way of broadening one's horizon, an enterprise that can succeed only if linguistic novelties are fed back into the familiar conceptual framework such that they become cognitively accessible.⁶⁸ Herder's account of translation thus reveals that an engagement with new manners of thinking is desirable only if this does not lead to forlorn contemplation of the unfathomable, but instead results in the attempt to render comprehensible the newly encountered by integrating it into one's prevalent cognitive structures. Imagination here aims for an alteration, refinement and enrichment of one's cognitive grasp of the world rather than a revolution of one's existent beliefs.

⁶⁷ It is interesting to note in this context that Heinz describes the process through which Herder's subjects acquire knowledge as an assimilation of certain aspects of the world; see Heinz (1994, 155).

⁶⁸ *Fragmente, zweite Sammlung*, 199–207, F 37–44. Cf. note 48.

8.5 Conclusion

By examining Herder's holistic conception of the mind in this essay, I have tried to show that in order to understand what is at stake in the eighteenth-century attack on metaphysical speculation we must distinguish between two forms of imagination. One that is experience-focused and conducive to the expansion of knowledge in its capacity to present to the mind particular aspects of reality, and a less desirable form that dovetails with a practice of reasoning that fails to remain in touch with the mind's sensibilities. Since imagination in the first sense of the term builds a connection between the mind and the world as it presents itself in experience, its integration becomes necessary for reasoned thought able to capture the reality of our lives. When the imagination is integrated in this way, it enables the kind of philosophy that many eighteenth-century thinkers advocated as an alternative to speculative metaphysics.

With this in mind, it finally becomes clear that the demand for a world-focused and experience-grounded philosophy is conceptually tied to the demand to think of reason in such a way that contributions of pre-rational elements to cognitive processes can be appreciated. Formulated like this, the polemics against metaphysics no longer seem to signify a move towards a simplistic glorification of observation and verification, but rather represent the attempt to complicate our understanding of how cognition works.⁶⁹

References

- Bacon, Francis. 2002. In *The New Organon*, ed. Lisa Jardine and Michael Silverthorne. Cambridge: Cambridge University Press.
- Beiser, Frederick C. 1987. *The fate of reason: German philosophy from Kant to Fichte*. Cambridge, MA: Harvard University Press.
- Buffon, Georges. 1750. *Histoire naturelle, générale et particulière*. 3 vols. Paris: LaHaye (First published in 1749).
- Condillac, Etienne Bonnot de. 1746. *Essai sur l'origine des connaissances humaines*. Amsterdam: P. Mortier.
- Condillac, Etienne Bonnot de. 1780. La logique, ou les premiers développemens de l'art de penser. In *Oeuvres complètes*, ed. G. Arnoux and Gabriel Bonnot, abbé de Mably. 23 vols. Paris: Houel.
- Condillac, Etienne Bonnot de. 1991. *Traité des systèmes 1749*. Paris: Fayard.
- D'Alembert, Jean Le Rond. 1995. *Preliminary discourse to the encyclopedia of Diderot 1751*. Chicago: University of Chicago Press.
- DeSouza, Nigel. 2012a. Leibniz in the eighteenth century: Herder's critical reflections on *The principles of nature and grace*. *British Journal for the History of Philosophy* 20: 773–795.

⁶⁹ I am grateful to Nigel DeSouza and Stephen Gaukroger for helpful discussion. I also want to thank the audiences of the "Nature versus Normativity" workshop in Berlin and the "Representation and Sensibility in Early Modern Philosophy" conference in Sydney in 2011 for their comments and questions.

- DeSouza, Nigel. 2012b. Language, reason and sociability: Herder's critique of Rousseau. *Intellectual History Review* 22: 221–240.
- Gaukroger, Stephen. 2001. *Francis Bacon and the transformation of early-modern philosophy*. Cambridge: Cambridge University Press.
- Gaukroger, Stephen. 2010. *The collapse of mechanism and the rise of sensibility*. Oxford: Clarendon.
- Haller, Albrecht von. 1757. *Elementa physiologiae corporis humani*. 8 vols. Lausanne: Marc-Michael Bousquet.
- Haller, Albrecht von. 1981. On the sensibility of irritable parts of animals 1755. In *The natural philosophy of Albrecht von Haller*, ed. Sirley Roe, 651–699, i–xxxii. New York: Arno.
- Heinz, Marion. 1994. *Sensualistischer Idealismus: Untersuchungen zur Erkenntnistheorie und Metaphysik des jungen Herder (1763–1778)*. Hamburg: Meiner.
- Herder, Johann Gottfried. 1985. *Werke in zehn Bänden*. Frankfurt: Deutscher Klassiker Verlag.
- Herder, Johann Gottfried. 2002. *Philosophical writings*. ed. and Trans. Michael Forster. Cambridge: Cambridge University Press.
- Kant, Immanuel. 1902. *Kants Werke, Akademie Ausgabe*, vol. 8. Berlin: Walter de Gruyter.
- Locke, John. 1975. In *An essay concerning human understanding 1700*, ed. P.H. Nidditch. Oxford: Clarendon.
- Longenesse, Beatrice. 1998. *Kant and the capacity to judge: Sensibility and discursivity in the transcendental analytic of the critique of pure reason*. Princeton: Princeton University Press.
- Nevinson, Henry. 1884. *A sketch of Herder and his times*. London: Chapman and Hall.
- Reich, K. 2001. In *Gesammelte Schriften*, ed. Manfred Baum. Hamburg: Meiner.
- Riskin, Jessica. 2002. *Science in the age of sensibility*. Chicago: University of Chicago Press.
- Sauter, Eugen. 1910. *Herder and Buffon*. Rixheim: F. Sutter.
- Schneiders, Werner. 1985. Der Philosophiebegriff des philosophischen Zeitalters: Wandlungen im Selbstverständnis der Philosophie von Leibniz bis Kant. In *Wissenschaft im Zeitalter der Aufklärung*, ed. Rudolf Vierhaus, 58–92. Göttingen: Vandenhoeck & Ruprecht.
- Sikka, Sonia. 2007. Herder's critique of pure reason. *The Review of Metaphysics* 61: 31–50.
- Spitz, Lewis. 1955. Natural law and the theory of history in Herder. *Journal of the History of Ideas* 16: 453–475.
- Taylor, Charles. 1991. The importance of Herder. In *Isaiah Berlin: A celebration*, ed. Edna Ullmann-Margalit and Avishai Margalit, 40–63. Chicago: University of Chicago Press.
- Vanzo, Alberto. Experimental philosophy in eighteenth-century Germany (Manuscript).
- Waldow, Anik. 2010. The pretense of skepticism and its nonepistemological relevance in early modern philosophy. *History of Philosophy Quarterly* 27: 35–56.
- Wenzel, Christian Helmut. 2001. Beauty, genius, and mathematics: Why did Kant change his mind? *History of Philosophy Quarterly* 18: 415–432.
- Wolff, Christian. 1741. *Elementa matheseos universae*. Halle: Renger.
- Wolff, Christian. 1755. Praefatio Physicae Electivae Jo. Christoph. Sturmii Tomi II. In *Meletemata mathematico-philosophica*. Halle: Renger.
- Zammito, John H. 2002. *Kant Herder, and the birth of anthropology*. Chicago: Chicago University Press.
- Zeuch, Ulrike. 1998. *Sentio, ergo sum*. *Herder Yearbook* 4: 143–155.

Chapter 9

Extending Nature: Rousseau on the Cultivation of Moral Sensibility

Annette Pierdziwol

9.1 Introduction

In an oft-quoted fragment Kant offers the following acknowledgment of the profound impact Rousseau had upon his way of perceiving others. He writes:

By inclination I am an inquirer. I feel a consuming thirst for knowledge, the unrest which goes with desire to progress in it, and satisfaction in every advance of it. There was a time when I believed this constituted the honor of humanity, and I despised the people, who know nothing. Rousseau corrected me in this. This blinding prejudice disappeared and I learned to honor man.¹

While this comment is often considered in relation to the idea of equality, it also seems to contain – perhaps unwittingly but nonetheless rather insightfully – a host of other motifs that gesture towards central preoccupations of Rousseau’s oeuvre. In this (albeit very brief) narrative Kant depicts himself as previously suffering from a “blinding prejudice” that decisively shaped his very way of seeing and feeling his relationships with his fellows. Interestingly, he describes this prejudicial mode of sight as a function of both cognitive ideas and of powerful affective feelings – his inclinations, thirsts and pleasures are all tied up in it. He thereby indicates the deep purchase it had on his whole psychology: it was not just a belief he held that caused some erroneous ideas but something central to his conception of himself (including his role as a philosopher) and his relation to others; something that deeply colored

I would like to thank Anik Waldow, Stephen Gaukroger and participants at the 2011 University of Sydney workshop on ‘Representation and Sensibility’ for their feedback on earlier versions of this chapter.

¹ Kant quoted in Dent (2005, v).

A. Pierdziwol (✉)
Institute for Advanced Studies in the Humanities,
University of Edinburgh, Edinburgh, UK
e-mail: annette.pierdziwol@sydney.edu.au

his moral perception. Furthermore, despite the stubborn hold this ‘despising’ way of perceiving others had on him, Kant testifies that it was corrected and that he was able to learn to relate to ‘man’ anew.

Read this way, Kant’s brief homage seems to have its finger on the pulse of a few of Rousseau’s greatest concerns, chief among these being the problems posed by such pervasive “blinding prejudice” – those deep-seated distortions of our moral perception – and so the pressing question regarding whether and, if so, how they can be corrected. Rousseau, of course, is notorious for finding fault with his contemporary society in such terms. He can frequently be found diagnosing the majority of his fellows as suffering from deeply mistaken conceptions regarding their needs, vulnerabilities, relations to others and so on. He depicts them as living under a fundamental error or illusion, which distorts – or, in one of his more potent images, poisons – the entire way they see, feel, reason, imagine, desire and act.²

Much has been written concerning the content of this error, drawing in particular on Rousseau’s discussions of amour-propre. In summary, its key feature is that it is a way of seeing and feeling one’s relations with others that operates via comparison, thus generating competitive relationships in which one is obsessed with one’s position of relative superiority and ends up, as Kant had it, habitually despising others.³ To perceive and feel one’s relationships with others in such comparative terms, according to Rousseau, is to be “wrong” and “deceived”⁴; it is to have a profoundly distorted moral sensibility. However, rather than debating the specific content of this error, my focus in this chapter will be on Rousseau’s insightful investigations into the *nature* of distortions of this type. In particular, I will say a lot more about his analysis of the psychological breadth and depth – hence, the pervasiveness and stubbornness – of the sort of mistakenness he has in view. His key insight here will be shown to consist in the way he analyzes its grip not just as an intellectual matter but as pertaining to one’s entire “way of being, of seeing, and of feeling”⁵ – thus fleshing out the significance of that affective dimension implicit in Kant’s comments.

Yet, further to this, while Kant’s brief testimony retells his overcoming of moral blindness as a seemingly instantaneous disappearance of prejudice and gaining of clear sight, it is on the question of the preconditions and mechanics of this process of transformation itself that Rousseau’s careful empirical investigations so often focus. Moreover, these investigations bring to the fore serious difficulties that in

²For example, the language of error and illusion can be seen in Rousseau’s descriptions of them as having imaginations ‘full of countless vain projects, troubled by countless whims’, as desiring ‘chimeral goods’ and as having minds filled with ‘countless ridiculous prejudices’ (Rousseau 2010, 409).

³Rousseau himself speaks in similar terms, for example in the *Emile* when describing the temptations with which Emile might be faced, he writes that in considering his rank he will be tempted to “say to himself, “I am wise, and men are mad.” In pitying them, he will *despise* them; in congratulating himself, he will esteem himself more, and in feeling himself to be happier than them, he will believe himself worthier to be so. This is the *error* most to be feared, because it is the most difficult to destroy” (Ibid., 400, my emphasis).

⁴Ibid., 401.

⁵Ibid.

fact make him rather pessimistic about the possibility of such correction at all, especially for adults whose distorted ways of seeing and relating to others already constitute deeply-ingrained habits. Recognizing this, Rousseau's own approach to questions concerning the possibility and paths of correction focuses enormous attention on seeking to understand the *source* of these distortions in the first place: how they arise and take on their characteristic shape.

Rousseau's answer in short, which I explore over the course of the chapter, is that these ways of seeing and feeling are a function of a person's whole psychology and that this in turn, as Rorty sums it up, "does not depend wholly on us as individuals: it depends on our early education and our social environment."⁶ Working largely within the framework of a sensationalist epistemology, Rousseau will argue that everything we are surrounded by – all the impressions to which we are exposed – gradually form our capacities in distinctive ways (as well as the coordination and "power politics" operative between them),⁷ which together go to make up our habitual ways of seeing and being affected in our relations with others.

Rousseau also frequently articulates this point in terms of sensibility, speaking, for example, of the way various impressions can impact the strength and trajectory of our moral sensibility, which he defines as "nothing other than the faculty of attaching our affections to beings who are foreign to us."⁸ For Rousseau there are really only two possible directions here, as he explains it, the strength of moral sensibility "is in proportion to the relationships we feel between ourselves and other beings, and depending on the nature of these relationships it sometimes acts positively by attraction, sometimes negatively by repulsion, like the poles of a magnet."⁹ Thus, for Rousseau, moral sensibility can be influenced to take a positive (extending outwards, attracted to others) course or a negative (contracting inwards, repelled from others) one. It is clear then that what Rousseau denotes by moral sensibility tracks closely with our way of perceiving and feeling our relations with others. What speaking in terms of sensibility furthermore allows him to capture is that what is at stake here is not various discreet instances of moral perception but rather something deeper and more pervasive, namely, as he puts it with regard to Emile, it is the orientation of his heart, the bent of his character or the "habitual condition of [his] soul" and the passions that dominate it.¹⁰ This is what is at stake in investigating the educational means by which moral sensibility is cultivated.

Before proceeding to these explorations, however, another issue raises its head here and along with it potentially a whole host of complications. This issue stems from the fact that anyone familiar with Rousseau's elucidations of the illusions, distortions and falsity characteristic of his contemporaries' way of relating to others will know that he typically cashes this out in terms of an appeal to nature.

⁶ Rorty (1998, 238).

⁷ Ibid., 237.

⁸ Rousseau quoted in Vila (1998, 184).

⁹ Ibid.

¹⁰ Rousseau (2010, 372, 373, 383, 389).

As Genevieve Lloyd puts it in the case of false ideas, “closeness to Nature is the mark of what is true in Reason, as opposed to what is false and factitious. Reason is held in check by Nature; it is not allowed to escape into the distorting realm of abstract speculation.”¹¹ Likewise for sensibility, imagination and desire: their ability to remain free from illusion would hinge on their naturalness, while their falsity would be a function of their artificiality. For Rousseau then, that way of seeing and feeling which will be characterized by perceptual lucidity is the one that is in sync with nature.¹²

It thus seems evident that Rousseau wants to invoke naturalness in some sense as the criterion for judging the satisfactoriness of various ways of life, i.e., that he invokes nature as a normative standard. It is the measure by which he deems his contemporaries mistaken, prejudiced, blinded and either insensible or hyper-sensible in their relations to the world and others. Yet, Rousseau’s appeal to nature as a normative standard is no small topic and has been subject to much debate. It is worth investigating this question then before proceeding. In particular, here we might worry that any analyses Rousseau has to offer regarding “blinding prejudice” will just end up in a simplistic recourse to some ideal of the natural. I will first briefly outline why this won’t be the case for Rousseau by showing how his obsessive interest in the truth of our natures in fact spurs rather than short-circuits his investigations into the falsity and distortion that frequently characterize our ways of seeing and feeling the relationships we have with others.

Section 9.3 then moves on to examine Rousseau’s ideas about education in more detail, firstly as they emerge in his perceptive analysis of the way society educates us into a distinctive – and, in Rousseau’s opinion, profoundly distorted – mode of moral perception. While Rousseau’s diatribes on this topic are well known, here I focus on his insight into what accounts for the terrible success of society’s educational method, namely, the way in which over time, through what it exposes us to and surrounds us with, it trains us in certain emotional, imaginative and affective habits as much as imparting certain ideas. In Sect. 9.4 I then consider how Rousseau seeks in the *Emile* to offer a sentimental education to rival this by outlining a comprehensive, pragmatic program in which Emile is given bodily practice at seeing, feeling, imaginatively identifying with and responding beneficently to real suffering. Finally, I explore how Rousseau’s constant emphasis on the difficulties involved in managing any such program bring to the fore a view of the cultivation of moral sensibility as a complex art and life-long task.

It is then perhaps Kant’s mention in the opening quote of his *learning* to honor man – implying, we might conjecture, an admission that he was not just able to see anew all of a sudden but had to *learn* how to do so – that is perhaps the most Rousseauian moment of his tribute. As will be seen, for Rousseau, it is ultimately only via the gradual process of a rival experiential education that one might cultivate another way of seeing and feeling one’s relationships with others.

¹¹ Lloyd (1983, 321).

¹² The term “perceptual lucidity” comes from O’Neal (1985, 56).

9.2 Unnatural Distortions

One of the things for which Rousseau is infamous is his emphatic denunciation of the artificial, alienating corruptions of society and his equally emphatic – even rhapsodic – celebrations of nature and “the natural man”. However, what becomes less clear once one begins to engage with the complexities of his accounts of nature and society beyond any simple polarity is how to read this appeal to nature. In his own time, some critics saw in this a rigid dichotomy, thus implying a ludicrously utopian call to withdraw from civilization and pursue a return to pre-social man as he might have been in the state of nature.¹³ However, the vast majority of Rousseau’s readers have resoundingly rejected such a construal of his chief enterprise, arguing that it is evidently not the case that he thought it was possible or even desirable to try make our way back to some supposed natural origin prior to all artifice.¹⁴

Instead, scholars typically maintain that whatever Rousseau’s obsession with nature, it is employed in the service of an obsession with the problems and possibilities of current social life: it functions as a diagnostic tool, a means for generating debate about the adequacy of our contemporary moral and social lives.¹⁵ Rorty, for example, argues that Rousseau’s interest in nature is aimed at providing “the grounds for evaluating social and political systems, and for diagnosing their failures” and thus is “also meant to indicate the moments for educational intervention.”¹⁶ Or, as Scott argues more directly, “his portrait of nature and of human nature is useful not as a focus of nostalgia but as a model for the conditions of our happiness, in accordance with which it is possible to manipulate the determinants of our nature.”¹⁷ In short, most critics seem agreed that Rousseau invokes nature not as something opposed to society, hence suggesting that getting close to the former would necessarily involve withdrawing from the latter, but rather that he invokes it

¹³ As Voltaire summed up this response in a rather nasty letter: “The desire to walk on all fours seizes one when one reads your work. However, as I lost that habit more than sixty years ago, I unfortunately sense the *impossibility* of going back to it, and I abandon that natural gait to those who are worthier of it than you and I” (Voltaire quoted in Johnston 1999, 30).

¹⁴ Dent (2005, 41). For example, Kant, for one, recognized this of Rousseau, as Cassier explains: “Kant judges that Rousseau’s purpose did not involve inviting man to go back to the state of nature but rather to look back to it in order to become aware of the errors and weaknesses of conventional society” (Cassier quoted in Marks (2005, 114).

¹⁵ In contrast then to the way Locke or Hobbes appeal to nature as a negative standard, it seems evident that Rousseau wants to invoke it positively as a normative guide of some sort for the sake of evaluating and improving current arrangements (Marks 2005, 17). As Dent puts it, he wants to use “the criterion of ‘naturalness’ as a test and measure of the satisfactoriness or otherwise of various particular social attitudes, relations, institutions” (Dent 1988, 16). Marks too offers “a Rousseau for whom nature remains a guide and limit for human beings in their pursuit of the best way of life” (Marks 2005, 118).

¹⁶ Rorty (1998, 239).

¹⁷ Scott (2006, 246).

as a norm for present social life. In other words, they argue that the decisive polarity of his thought is not between a pre-social nature and society but is rather between an unnatural sociality and a natural sociality.¹⁸

The question, however, is whether Rousseau can in fact invoke nature in this way and what, if anything, it amounts to in the end. Judging this depends largely on what one makes of Rousseau's account of nature. Here, though, any scholarly consensus fast begins to come unstuck. This is partly due to the fact that Rousseau's writings themselves contain at least two seemingly contradictory definitions of nature between the *Second Discourse* and later works like the *Emile*.¹⁹ This has allowed room, roughly speaking, for two different camps of interpretation on these questions to emerge. The first is a product of those who tend to focus on the former text. They insist that here we encounter a Rousseau who has a decidedly modern, minimalist definition of nature in line with the scientific aspirations of his age: nature refers to origins.

For those who take this view, if it is the case that Rousseau appeals to nature as a normative standard then it is one that has been emptied of all content. For some this is bad news, as it would seem to spell doom for nature's ability to offer any guidance. For others, however, this is its great (and perfectly Kantian) promise: Rousseau invokes nature, only to then reveal its emptiness. In doing so, he demonstrates the need to put in its place an appeal to freedom or perfectibility. Thus, these critics argue, when we actually look closely at Rousseau's definition of nature – its peculiar emptiness and so inability to provide any standard – what we discover is that Rousseau's invocations of nature as a normative standard undergo a metamorphosis into something quite different altogether. As Marks nicely sums it up, the conclusion here is that "Rousseau returned to nature only to murder it."²⁰

However, another camp of interpretation on Rousseau's account of nature also exists. This approach, largely drawing on the *Emile* argues that for Rousseau natural man is not simply equated with original or prehistoric man prior to the emergence of society. Rather, here we see quite plainly that Rousseau has a much more capacious understanding of nature. This can be seen firstly in Rousseau's explicit clarifications in this text that when he declares that his goal is "to form the man of nature" he does not have in mind some naïve return to "the depths of the woods"²¹ – some attempt to recreate "the natural man living in the state of nature"²² by removing the child at birth from all social artifice and allowing nature to unfold with no interference whatsoever (the usual caricature of negative education). In fact, as many commentators have pointed out, *Emile*'s education is a thoroughly social project, full of its own peculiar brand of artifice and, moreover, as Rousseau explicitly states, its aim is to make *Emile* "the natural man living in the state of society"; "a savage made to inhabit cities."²³

¹⁸ Cooper (1999, 48–50).

¹⁹ For further on this point, see Marks (2005, 111).

²⁰ *Ibid.*, 20.

²¹ Rousseau (2010, 412).

²² *Ibid.*, 355.

²³ *Ibid.*

In the text we see that this involves a concerted attempt to make Emile into all he can be: to cultivate, train and teach him in various ways that will extend and strengthen the “original dispositions born from the use of his senses.”²⁴ Thus, just as Rousseau celebrates the way the blind have learnt “a surer and keener touch”²⁵ than that which we typically possess (hence testifying to the plasticity of the senses and their potential for extension), so too Rousseau encourages Emile to undertake practices to stretch his capacities – to “get from each of them all that they can do.”²⁶ Indeed, his ambition here frequently soars well beyond what might be expected from letting nature (our original dispositions and nascent sensibilities) unfold unassisted.²⁷ Likewise with regard to the moral realm, Rousseau is certainly keen to cultivate all the ‘higher’ capacities in Emile: “sublime sentiments,” “judicial clarity,”²⁸ justice, goodness, humanity, commiseration and beneficence.²⁹

Very significantly, it is this final portrait of a fully cultivated Emile that is Rousseau’s ‘natural man’ – his vision of natural sociality or, in Cooper’s terms, “civilized naturalness”³⁰ – and so it seems that Rousseau’s understanding of nature in fact involves a “peculiar blend of ‘high’ and ‘low,’ or ‘moral’ and ‘physical’.”³¹ In the *Emile* Rousseau makes it very clear that the products of such cultivations or educational extensions of one’s native endowment, which are achieved “as we become more capable of using our senses and more enlightened,” are in fact to be considered part of nature, as he puts it, they are “what I call in us *nature*.”³² Here then nature certainly does not equal pre-social origins. Rather, the natural man – just like the unnatural man – is the product of countless cultivations of his nature.³³ What makes these “extensions” natural, however, is that they proceed in a manner that is in accordance with – i.e., one that strengthens and deepens rather than corrupts – those original dispositions.³⁴ For Rousseau then the name nature denotes “habits conformable to nature.”³⁵

But it might immediately be asked how we can judge this conformity. How do we know what counts as an extending and strengthening of nature and what as a corruption of it? Are we not just returned again to debates over the meaning of this original nature for Rousseau and whether it in fact has any normative potential – whether as

²⁴ *Ibid.*, 163.

²⁵ *Ibid.*, 273.

²⁶ *Ibid.*, 272.

²⁷ For further on this, see Marks (2005, 45).

²⁸ Rousseau (2010, 410).

²⁹ *Ibid.*, 389, 375.

³⁰ Cooper (1999, 10–11).

³¹ *Ibid.*, 69.

³² Rousseau (2010, 163).

³³ *Ibid.*, 387.

³⁴ Cooper (1999, 64).

³⁵ Rousseau (2010, 163).

a formal model or more substantive content³⁶ – to guide us in our cultivating enterprises? Some contend that this certainly seems to be the case. However, a distinctive feature of the arguments of those commentators who take the *Emile*'s more capacious definition of nature seriously is that most end up highlighting the way in which, via this definition of nature, “the whole mode of describing and investigating nature takes on a historical or developmental character.”³⁷ And in so doing, serves to call into question the idea that we might be able to define nature in advance apart from this development.³⁸ Rather, since natural man on the *Emile*'s definition is the outcome of a developmental process, it seems we will only be able to work out what counts as natural via engaging in such processes and examining their end products.³⁹ In other words, establishing what “the contours of our true nature are and what way or ways of life are compatible with it”⁴⁰ can only be worked out via the close study of various processes of ‘extension’ or cultivation. This is precisely what Rousseau seeks to investigate with regard to his contemporaries and it is also the point of his philosophical educational experiment in the *Emile*.⁴¹ He claims that in order to judge on the question of naturalness, the natural man

would have to be seen wholly formed: his inclinations would have to have been observed, his progress seen, his development followed. In a word, the natural man would have to be known. I believe that one will have made a few steps in these researches when one has read this writing.⁴²

Returning then to the question of the role nature might play in Rousseau's evaluations of the perceptual lucidity of various ways of seeing and feeling our relationships with others, it now seems evident that this will be a complicated affair. Indeed, it seems that the appeal to nature far from short-circuiting his inquiries with a pre-given set of answers about what would count as natural, instead gives rise to a certain form of inquiry – of educational experimentation and observation – which can only flesh out an answer to this question concerning the ways of life compatible with nature and the signs by which they can be recognized amongst the complexities

³⁶ Scott (2006) offers arguments in support of a formal model, while Cooper (1999) argues for a more substantive content.

³⁷ Marks (2005, 113).

³⁸ Lloyd also captures this idea in her claim that “Reason and Nature are not, for Rousseau, equal and independent terms, complete in themselves; ...their interdependence is quite complex” (Lloyd 1983, 323).

³⁹ Thus, as Grimsley puts it, “the true significance of nature will appear only gradually” (Grimsley 1983, 48).

⁴⁰ Marks (2005, 51).

⁴¹ Here commentators also point to other texts, particularly Rousseau's autobiographical works, which provide a similar type of close “study and description of extraordinary types and their way or ways of life” (Marks 2005, 51).

⁴² Rousseau (2010, 165–6).

of a developmental account.⁴³ And, as noted, it is precisely these lengthy processes of experience, analysis, comparison and so on to which Rousseau invites us in works like the *Emile*. In short then, far from yielding a ready-made set of “pat answers” and “simple solutions,”⁴⁴ Rousseau’s appeal to nature as the measure of the truth of our cognitive, emotive and moral relation to the world – of the perceptual lucidity or distortedness of the way we see and feel – is what gives his investigations their highly distinctive set of questions and methodology: it is no accident that the *Emile* takes the kind of developmental-novelistic form it does.⁴⁵

Of course, one obviously significant path of inquiry with regard to an account of this type would be to look at its end-products: to assess Rousseau’s final vision of natural sociality as exhibited in his fully formed “natural men”, such as Emile, in order to ascertain the specific criteria of naturalness at which he eventually arrives via this approach. Thus, for example, with regard to moral life, which is the focus of the present chapter, we could study Emile’s developmental trajectory and final character in order to establish what Rousseau thinks a natural, undistorted way of seeing and feeling one’s relations with others might involve. We could also supplement this by comparing it with Rousseau’s descriptions of fully formed ‘unnatural men’ in order to again further specify the criteria he arrives at to define the artificiality constitutive of their lives and relationships.

This avenue of critical analysis would thus certainly be vital to assessing the overall success of Rousseau’s project and, indeed, much scholarly debate has focused on discerning these criteria and debating their ultimate convincingness.⁴⁶ With regard to the moral sphere, as noted in the introduction, the key content of this criterion of unnaturalness in one’s relations with others is usually specified in terms of their comparative nature, “the habit of measuring oneself against others,”⁴⁷ of

⁴³ Cooper also highlights the way the appeal to nature does not provide any kind of easy short-cuts for Rousseau: “But this guidance [of nature] is neither easy to attain nor easy to follow. It is not a simple thing to extend and strengthen one’s natural dispositions or even to know how to try to do so. Once one has left the state of nature for the civil state, nature ceases to speak very clearly – or, which is much the same thing in its effect, one ceases to hear its voice very distinctly” (Cooper 1999, 64); its “guidance is much less direct and articulate” (Ibid., 9). As such, Cooper argues that this implies a new conception of the *work* we might need to do in order to access it. As he explains in a quote which highlights the crucially *experimental* flavor of Rousseau’s approach: If one wishes to find a standard in nature “he must be more creative. Rather than expect answers to his questions, *he must develop proposals of his own and then test them against nature*. Nature remains the final arbiter, but it merely nods, as it were, rather than speaks. No longer a source of positive guidance, it is at most a touchstone” (ibid., xiv, my emphasis). As such, as already suggested, the task of discerning what counts as natural in society will be a lengthy, complex process that can only be worked out developmentally in experience.

⁴⁴ Marks (2005, 116, 117).

⁴⁵ For further on this point, see Cooper (1999, 68–9).

⁴⁶ For example, chapter two of Cooper (1999) “aims to ascertain the criteria of post-state-of-nature naturalness” (Cooper 1999, 12). However, final conclusions regarding these criteria are seldom entirely favorable especially when the further developmental stages involving Sophie are taken into account. Thus, for example, Marks concludes that in the end we are left to wonder “if Rousseau’s account of nature’s end is not misleading and exaggerated” (Marks 2005, 116).

⁴⁷ Rousseau (1990, 112).

perceiving and feeling things in terms of one's difference from others. By contrast, natural sociality – that undistorted moral sensibility being championed for Emile – seems to consist principally in his perceiving others in terms of their identity with him as members of the same species, all ultimately “subject to the same weaknesses”⁴⁸ and so sharing “the same passions, the same sentiments.”⁴⁹ Much more could be said and debated about this criterion, for example, about whether Rousseau provides sufficient argument or evidence for why exactly comparative ways of seeing and feeling are the key problem, or, perhaps one might want to contest Rousseau's highly unflattering descriptive portrait of his ‘unnatural men’ by offering a rival developmental story that embodies a different vision of natural sociality.⁵⁰

Yet while these are crucial debates to continue for the critical assessment of Rousseau's conclusions, for the purposes of the present chapter I want to attempt to step somewhat to the side of questions concerning the specific content of Rousseau's criterion of naturalness, as exhibited in the end products of his educational experiments, in order to instead focus on what emerges through the process of these analyses themselves. In particular, to begin with in the next section I seek to show how through the course of these investigations Rousseau offers a penetrating analysis of the *type* of problem at stake in the unnatural distortions of moral sensibility with which he is concerned. As a result, even if we end up disagreeing with him about the content of that distortion (i.e., even if we find his ultimate portrait and criterion of the “natural man” and his mode of moral perception deeply unconvincing), we may nonetheless still find along the way of his analysis much food for thought concerning the nature of such deep-seated distortions – insights that serve to foreground further significant questions regarding their source and possible means for overcoming them. The remainder of this chapter then will attempt to show the way in which Rousseau's engagement with these questions might make his writings instructive even to those who cringe at the appeals to nature which motivate them.

9.3 Society's Education

Everywhere there are excesses to fear.⁵¹

As already briefly outlined, Rousseau thinks there is something profoundly wrong with his contemporaries' moral sensibilities. He is deeply concerned about

⁴⁸ Rousseau (2010, 401).

⁴⁹ *Ibid.*, 377. Note that given there is no putting off the birth of amour-propre, Rousseau will go on to argue that it is also necessary to school Emile in the inequality that characterizes society but this ability to see people in terms of their differences comes later and concerns Emile's political rather than moral development, see Rousseau (2010, 389).

⁵⁰ This would be to engage Rousseau on his own methodological terms since as suggested above for him what is ‘natural’ is discerned only in and through these processes of formation and not at some pure origin prior to sociality. As such, an alternative criterion of naturalness (natural sociality) could only be gained by telling a different developmental story.

⁵¹ Rousseau (2010, 384).

the falsity or distortedness of the way they see, feel, desire, imagine, judge and respond – in short, with their entire “way of being”⁵² – which he often casts in terms of its unnaturalness. As has been seen above, however, a number of very significant complexities attend this invocation of nature. Furthermore, as also already noted, my aim here is not to enter into debates about the specific content and validity of this diagnosis of artificiality. Rather, my interest lies in tracing the way in which Rousseau’s obsession with “the man of nature” (i.e., with envisioning a natural sociality) and his method (i.e., as argued above the necessarily developmental approach his conception of nature requires he take) lead him in the process to offer highly insightful analyses concerning the *type* of problem at stake.

I have already gestured towards the central element in this analysis in the introduction, namely, that Rousseau is concerned with a problem that involves all one’s cognitive, affective and imaginative capacities and so one that is tied up with one’s entire psychology at a fundamental level. Building on this, it seems that Rousseau presents this problem as global in its scope, by which I mean he depicts it not simply as having to do with errors that crop up on various occasions but rather as involving a kind of *systematic* mistakenness that fundamentally skews one’s entire sensory-affective experience. Rousseau thus often draws attention to the distorting selectivity of what he deems the artificial, comparative mode of moral perception and the way it serves in advance to fundamentally structure what shows up as cognitively and affectively salient for us, potentially creating certain blind spots or nodes of insensibility:

[*Amour-propre*] is irritated by the advantages someone else has over us, without being appeased by those for which it feels compensated. The feeling of inferiority is a single respect *poisons* the feeling of superiority in a thousand others, and what one has more of is forgotten in devoting attention only to what one has less of.⁵³

Interestingly, the language of sensibility, with which Rousseau and many other French Enlightenment thinkers work, lends itself remarkably well to formulating observations of this sort. The notion of a disordered sensibility (whether of one’s being insensible or having an excessive sensibility for example) could capture the idea of there being a problem not merely with how a person interpreted or made judgments about their experience, but of there being a problem at the level of experience itself. That is, the idea that the very way in which a person sensed and was affected by things could be in some sense out of touch with the external world of things and others supposedly giving rise to these responses. Understood to be at issue then was a rather pervasive type of problem: something that *systematically*

⁵² *Ibid.*, 401.

⁵³ Rousseau (1990, 113) (my emphasis). A similar idea is found in the *Emile* where Rousseau writes, “And were he despised by only a single man, that man’s contempt instantly poisons the others’ applause” (Rousseau 2010, 381). Another example of Rousseau depicting artificial moral sensibility as involving a systematically distorted way of experiencing things is found where he writes: “Relating everything to themselves alone and regulating their ideas of good and bad according to their own interest, they fill their minds with countless ridiculous prejudices, and in everything that hampers their slightest advantage, they immediately see the overturning of the whole universe” (*ibid.*, 409).

distorted one's *entire* way of perceiving, being affected by and responding to others. It is no wonder that Rousseau also frequently depicts such ways of seeing as having the form of addictive habits, as he dramatically puts it on one occasion, "once we have started to measure ourselves this way, we never stop."⁵⁴ Indeed, given their role in defining the very way in which we experience things, Rousseau often seems doggedly pessimistic about our chances of breaking out of them at all. This analysis of our potential to suffer from certain systematic distortions of our moral perception (or, on Rousseau's diagnosis of his contemporaries, their actually suffering from them) thus seems to imply the near impossibility of our being able to remedy them in any quick or direct fashion (for example, by simply choosing to see and feel differently).

The significance then of Rousseau's elucidation of the global scope of this type of problem is that it directs him to focus considerable attention on seeking to understand the *source* of such systematic distortions in the first place: to trace the origins of how they arise and come to gain such a firm hold over us. His response to this question, which shares much in common with other French Enlightenment thinkers writing in the wake of sensationalism, is to argue that it is his contemporaries' social context and all the stimuli to which it exposes them that is educating or habituating them into these distorted moral sensibilities. To support this claim, Rousseau needs to tell a story about how this happens. In outline, his contention is that it is a function of the sheer excess of impressions with which people are surrounded in society.⁵⁵ This occurs, according to Rousseau, as the excessive multitude of objects and sentiments constantly striking and affecting us in the social whirlpool generate in us an exponentially-multiplying increase of needs and dependencies,⁵⁶ which he thinks inevitably lead us to glance sideways and inflame our imaginations with anxious curiosity about one another's opinions. In this way then the increase of objects, relations, ideas and so on (i.e., the excess of stimuli impressing upon one's sensibility) creates in us that habitual disposition to compare ourselves with others from which "all the hateful and irascible passions are born."⁵⁷

⁵⁴ Rousseau (1990, 113).

⁵⁵ But why should we think excess has this effect? Some hints can be found in the following quote where Rousseau contrasts earlier, simpler forms of society with the gradually increasing complexity that becomes characteristic of it: "As society becomes more closely knit by the bond of mutual needs, as the mind is extended, exercised, and enlightened, it becomes more active, embraces *more* objects, grasps *more* relationships, examines, compares. In these frequent comparisons, it doesn't forget either itself, its fellows, or the place it aspires to among them. Once we have started to measure ourselves this way, we never stop" (ibid., my emphasis).

⁵⁶ Rousseau (2010, 412).

⁵⁷ Rousseau (2010, 364). Rousseau explains how comparison breeds these passions as follows: "As soon as one adopts the habit of measuring oneself against others and moving outside oneself in order to assign oneself the first and best place, it is impossible not to develop an aversion for everything that surpasses us, everything that lowers our standing, everything that diminishes us, everything that by being something prevents us from being everything" (Rousseau 1990, 112). See also: Rousseau (2010, 364, 378).

Moreover, all of this is further exacerbated Rousseau thinks by the fact that contemporary Parisian society not only presents an excess of stimuli that generates comparative sentiments but that it also typically exposes its members to specific impressions that further entrench this, for example, ones he worries over in the *Emile* include luxuries, entertainments, constant scenes of health and happiness (i.e., the avoidance of all suffering), the pomp and splendor of rank, the appeal of riches, heights of glory, brilliant lots etc.⁵⁸ Rousseau seeks to map a connection between repeated exposure to such scenes and the development of certain habitual ways of thinking of oneself as different from and more important than others as well as habitual passions such as vanity, resentment or jealousy and habits of the imagination in which one finds it difficult to identify and empathize with the plight of others. While this is not the place to evaluate the evidence for these claims (though this is certainly the response they invite as purported empirical observations), the key point is that, although Rousseau thinks his contemporaries' distorted moral sensibilities indeed represent a very serious, pervasive type of problem, he also thinks it is nonetheless possible to uncover their origins in their social environment.

Of course, these points are by no means unique to Rousseau, as Natasha Gill notes, "in the area of pedagogy Rousseau was a disciple of the disciples of sensationist education theory,"⁵⁹ that is, of "the work of French pedagogical theorists as they interpreted and struggled with Locke's ideas,"⁶⁰ especially his epistemology's key premise of the child as a blank slate. One particularly significant point of continuity here would thus seem to lie in Rousseau's emphasis on psychological malleability, as Rorty puts it, on "our plastic susceptibility to social formation."⁶¹ On the French sensationalist account moreover this malleability was typically seen to go all the way down: theorists held that all the stimuli impressing via the senses not only give rise to the 'content' with which our various intellectual and affective capacities dealt but that it is these sensations which also serve to shape the development of those very capacities themselves: the decisive formations they take and the way they relate to one another.⁶²

Likewise, Rousseau seems interested in examining what effect society's excessive proliferation of stimuli is having on the formation of his contemporaries' capacities: what distinctive forms of imagination, rationality, desire, affectivity and so on is it

⁵⁸ Rousseau (2010, 373 ff).

⁵⁹ Gill (2010, 184).

⁶⁰ Ibid., 181. Gill, however, also goes on to locate profound points of discontinuity with this tradition, seeing Rousseau's *Emile* as designed to challenge and rethink many of its key tenets. I comment further on her arguments below.

⁶¹ Rorty (1998, 242).

⁶² As Gaukroger explains: "In his *Traité des sensations*, Condillac goes beyond Locke's enquiry into how ideas come into our mind, asking also about the origins of our mental faculties themselves" (Gaukroger 2010, 413–4).

molding?⁶³ In this way, the sensationalist account also allows him to explain why our ways of seeing and feeling, as noted above, typically seem to take the form of such addictive habits. For Rousseau, the stubbornness of our habitual modes of moral perception is a function of the fact that they stem from such deep-seated formations of our whole psychologies. Furthermore, as Rorty highlights, as society's sentimental education over time shapes a person's psychology in this way, he also ends up becoming "complicit in the process of his malformation."⁶⁴ His cognitive and affective capacities no longer offer any resources "to correct malformed, harmful passions,"⁶⁵ no means for breaking out of his distorted moral sensibility.

According to Rousseau then, his contemporaries' way of seeing and feeling and what he deems its characteristic distortedness is in fact the result of long processes of habituation and learning amongst society's excess of impressions. What might thus be termed 'society's education' consists not simply in its explicitly teaching certain cognitive ideas about our natures, needs, relations and so on but in the much more subtle operations by which it constantly exposes us to stimuli that tug on our senses, provoke our passions, shape desire and set the imagination racing. Rousseau's largely sensationalist pedagogical approach enables him to see the entire social milieu of eighteenth century Paris as an education: the excess to which it exposes its members for him amounts to nothing other than an intensive cultivation program – one which decisively forms all their capacities and so shapes the very way in which they perceive, are affected by and moved to action in their relations with others.

The problem Rousseau wants to highlight, however, is that for the most part his contemporaries seem oblivious to the way in which their context is training them in certain passionate and imaginative habits. Rousseau's aim is to open his readers' eyes to these unobtrusive but terribly efficient educational mechanisms: to the fact that their existence in society is not a neutral backdrop but a vast array of stimuli that impress upon them and profoundly shape the development of their cognitive and moral capacities, directing their sensibility along certain trajectories and bending the habitual passions of their souls into particular grooves.

Turning now to the implications of Rousseau's analysis, in the first instance it seems rather devastating in the way it confirms how deep the hold of such distortions will be, traceable as they are back to decisive formations of one's very psychology. On this picture, once the crucial stages in this "total formation of children's characters: their minds, their morals, their very nature"⁶⁶ has occurred (for Rousseau, adolescence is the key phase here), further reformation looks highly

⁶³ Some examples of Rousseau's attention to this question are provided by Rorty as she explains his account of how in society the imagination takes the form of "fantasy, still linked to the satisfaction of desire, but now constructing and exploring remote possibilities, whose conceptualisation generates more desires, and whose satisfaction creates even more refined possibilities" (Rorty 1998, 241). Likewise, rationality itself takes on a distinctive shape: "A calculating form of prudential rationality directed to satisfying desires develops in an uneasy relation to the fantasy-imagination" (ibid.).

⁶⁴ Ibid.

⁶⁵ Ibid., 242.

⁶⁶ Gill (2010, 7).

implausible. While, arguably, Rousseau never completely rules out the possibility of remedial work for adults whose moral sensibilities are already decisively shaped,⁶⁷ he certainly does not think this is where we should pin our hopes. Instead, like many of his fellow pedagogues, he focuses on the promise of early intervention in this educational process before its formations are able to become so stubbornly entrenched. This preference is evident as Rousseau describes the method for remedying prejudices potentially arising in Emile's moral perception:

One must undeceive him or, rather, *anticipate* the error for fear that afterward it will be too late to destroy it... For this there is no cure other than experience – if, indeed, anything can cure it. At its birth, at least, one can *prevent* its growth.⁶⁸

This then is the motivation for Emile's alternative education: to implement right from the start a rival schedule of exposure to impressions that might shape and form his cognitive, affective and moral capacities in a way quite different to that of society – a way that Rousseau hopes to convince his readers is far more in accord with nature; a strengthening and deepening of it. Against society's education then, Rousseau wants “to show what education can do for a man”⁶⁹ in forming him into “the man of nature”⁷⁰ (in producing a natural sociality). Nonetheless, as will become clearer in the next section, Rousseau's attentiveness to the implicit educational techniques society so effectively employs will offer vital lessons for his attempt to construct this rival program.

9.4 Cultivating Moral Sensibility

Unhappy men, dying ones, sights of pain and misery!⁷¹

In view of what has been said above, one way of formulating Rousseau's key question in the *Emile*, particularly at the start of Section IV where he is concerned with the commencement of Emile's moral development could be as follows: given that society, via all the stimuli with which it surrounds a person, would train Emile

⁶⁷ Rousseau suggests that if there is any hope with regard to even those most deeply-ingrained distortions of our moral perception, those that are “the most difficult to destroy” (Rousseau 2010, 400), then it will be the case that “[f]or this there is no cure other than experience” (ibid., 401). In this, we might see Rousseau as holding true to his experimentalist commitment to not rule out anything in advance of testing it out in experience. As he explains this principle in the context of stretching Emile's physical sensibility: “We can know the use of our organs only after having employed them. It is only *long experience* which teaches us to turn ourselves to account, and this experience is the true study to which we cannot apply ourselves too soon” (ibid., 289, my emphasis).

⁶⁸ Ibid., 401 (my emphasis).

⁶⁹ Ibid.

⁷⁰ Ibid., 412.

⁷¹ Ibid. 380.

in a certain distorted mode of seeing and feeling his relations with others, how might Emile's tutor not only inhibit this education but also construct an alternative program of cultivation?⁷² In short, how might Emile be educated into another way of seeing and feeling – one less prone to the unnatural distortions plaguing the comparative mode of perception? How can the tutor arrange things so as to enable him to acquire “a sense of the true relations of man”?⁷³ As Rousseau phrases it, what are the roads by which “it is fitting to penetrate the heart of a young adolescent in order to arouse the first emotions of nature and to develop his heart and extend it to his fellows”?⁷⁴ This will involve in the first place an attempt to stop society's education in its tracks but, importantly, it will also involve the attempt to construct a rival education that, as it turns out, is premised on a healthy respect for the former education's ingenious methods.

To begin with the first, most well known part: Rousseau, like many other Enlightenment worriers about the effects of an excessive milieu, starts with the need to moderate one's exposure to this excess of impressions. Thus in the *Emile* we find him frequently recommending that the tutor remove Emile from society or, if he must remain in it, that he be ever vigilant in controlling the impressions to which Emile is exposed: “Choose with care their society, their occupations, their pleasures.”⁷⁵ In making these claims, however, as noted in Sect. 9.2, Rousseau highlights that his aim is not “to make [Emile] a savage and relegate him to the depths of the woods” and yet, nonetheless, in this “social whirlpool” the tutor has to arrange

⁷² Gill (2010) argues that there is, however, a major tension surrounding this idea of education in Rousseau's work, which sets him quite apart from other French Enlightenment thinkers. Indeed, she goes so far as to suggest that he would reject the idea of offering an alternative “education” altogether if education means, as it did for most French sensationalists, mere external conditioning and habituation (Gill 2010, 201). Rousseau, Gill argues, is deeply concerned with the passivity and determinism this picture ends up producing, such that someone like Helvétius can write as if good citizens can be scientifically engineered (Gill 2010, 206). On Gill's account, Rousseau's *Emile* is specifically designed to buck this trend. Particularly though the story of the Savoyard Priest, she argues that Rousseau seeks to emphasize the role of our active capacities and some sort of inner moral sense (conscience) as vital to any moral development, hence rejecting the idea that virtue can be produced simply by orchestrating one's external environment. What makes this assessment complicated, however, is that, as Gill notes, Rousseau clearly never renounces the basic principles and methods of sensationalist philosophy: designing an alternative experiential education for Emile still remains the central plank of his approach. Thus, while outside the scope of this present chapter, Gill's work highlights the need to explore how the methods I will examine in this section relate to Rousseau's later claims in the Savoyard Priest passages. The tensions that remain in his account as a result of this (for example, his fascinating ambivalence about the role of habit in moral cultivation, see Gill (2010, 190 fn 12) are vital to understanding his attempt to fundamentally reinterpret the legacy of French sensationalist thought. Later in this chapter I will, however, broach one aspect of this conflict in Rousseau's “response to theorists' increasingly scientific approach to pedagogy” (Gill 2010, 187), which can be seen in his constant emphasis on any program of moral cultivation remaining a complicated *art*.

⁷³ Rousseau (2010, 371).

⁷⁴ Rousseau (2010, 378).

⁷⁵ *Ibid.*, 384.

things so as to help Emile “not to let himself get carried away by either the passions or the opinions of men.”⁷⁶ Limiting Emile’s exposure to the excess of objects on offer in eighteenth-century Paris – “the moderation of hearts”⁷⁷ – is thus central to the task of inhibiting society’s educational mechanisms and their impact on the formation of his capacities.⁷⁸

However, Rousseau also makes clear that this preventative path of moderation is only one side of the story for him. The aim is not just to keep Emile’s nascent sensibility on a leash in order to prevent the stimuli with which he is surrounded in society directing it along a certain negative course. For Rousseau, this approach could never be entirely successful unless the child is completely alone,⁷⁹ but then its success would also be its failure since “so long as his sensibility remains limited to his own individuality, there is nothing moral in his actions.”⁸⁰ Thus Rousseau often emphasizes that a complete withdrawal is plainly undesirable as far as Emile’s moral development is concerned. Rather, it is necessary that there is some increase of needs, relations and so on: the child’s nascent sensibility must develop and “extend outside of himself.”⁸¹ What is crucial for Rousseau is the direction of this extension.

In line then with what was argued in Sect. 9.2, for Rousseau, there is no question of a retreat to some pure original nature. Human social and moral development inevitably involves the artifice of cultivating and extending our original natural dispositions and this, as a rule, is a very good thing. Rousseau has no interest in stunting Emile’s growth by abandoning it to unfold without any artful intervention. His problem, recall, is not with cultivation per se but with the fact that he thinks society’s cultivation is out of sync with and a betrayal of nature. His own aim is to cultivate “the man of nature” (natural sociality), i.e., to extend and strengthen Emile’s original dispositions in a manner that he considers to be in accordance with nature. Thus while there will clearly be an emphasis on a “negative education” (removal from all social artifice) when it comes to inhibiting society’s cultivation program, it also seems that Rousseau places an equally strong emphasis on a positive education of sorts when it comes to the tutor’s own attempts to construct a rival educational plan that seeks – intentionally and with much artifice of its own – to extend nature along another course. As Rousseau himself confesses, it seems

⁷⁶ *Ibid.*, 412.

⁷⁷ *Ibid.*, 390.

⁷⁸ Gill also highlights this point: “Negative education is more than the elusive method Rousseau disarmingly describes as “doing nothing,” or “preventing anything from being done”... Behind the general advice “don’t do anything” lies a set of more specific rules that aim to inhibit particular educational practices advocated by many eighteenth-century theorists” (Gill 2010, 187).

⁷⁹ Rousseau (2010, 370).

⁸⁰ *Ibid.*, 371.

⁸¹ *Ibid.* Even where Rousseau admits there may also be no stopping the eventual birth of amour-propre, he again implies that the crucial question is how it is then directed, thus he admonishes: “Let us extend amour-propre to other beings. We shall transform it into a virtue” (*ibid.*, 409).

“one must use a great deal of art to prevent social man from being totally artificial.”⁸² In short, to reconnect with the conclusion of Sect. 9.2, this is again the point that, as Lloyd puts it, “nature needs education to be what it truly is.”⁸³

Once it is recognized then that Rousseau is not promoting a retreat to “the depths of the woods”⁸⁴ where the tutor lets ‘nature’ unfold completely unhindered but is rather interested in constructing an intentional program of cultivation to rival the one he sees implicitly at work in society, what can also be seen more clearly is that Rousseau’s own method for approaching this cultivation is in fact premised upon a healthy respect for the sheer effectiveness – the comprehensiveness and depth – of society’s educational methods. While he is obviously not a fan of its end products, Rousseau seems to imply that one must nonetheless seek to understand the secrets of its success. In particular, that what makes it so profoundly effective is that it surrounds us with stimuli that target the full gamut of our cognitive, affective and imaginative capacities, thus shaping the very way we see and feel. In the first place then Rousseau accepts that his own rival educational plan must work at a similarly comprehensive and deep level. This makes him impatiently dismissive of any approaches which attempt to educate one on “the true relations of man” by proceeding merely at an intellectual level. For Rousseau, there can be no real learning here which circumvents the body. The terrifying success of society’s jointly cognitive and affective education reveals the impotence of abstract, theoretical means. Rousseau’s goal is thus not only that Emile might gain knowledge about the true nature of his relations with others but that he “see with his eyes, that he feel with his heart.”⁸⁵ As such, his rival plan of sentimental education must also operate on all fronts if it is to be as deeply formative of Emile’s capacities and the trajectory of his moral sensibility as that of society’s.

How then does one go about designing an education of such breadth and depth? For Rousseau, still working largely within a sensationalist epistemological framework, this remains an inescapably physical issue: a matter of “carefully controlling the impressions” made upon one’s sensitive system.⁸⁶ In addition then to the moderated avoidance of excess required in order to inhibit the workings of society’s educational program, what emerges as equally important to cultivating moral sensibility is that the tutor exercise diligent selection over *which* objects and sights Emile is exposed to, i.e., that he deliberately put in place an alternative program of impression-exposure. Thus, the tutor must also “by the choice of circumstances in which [he]

⁸² Rousseau (2010, 485).

⁸³ Lloyd (1983, 325).

⁸⁴ Rousseau (2010, 412).

⁸⁵ Ibid.

⁸⁶ Vila (1998, 183). Commenting on Rousseau’s later declared intention to write a *morale sensitive*, Vila writes that Rousseau “believed that morality could be achieved through material means – that is, by a proper understanding and treatment of the body” (ibid., 186). This again, though, would sit in tension with the themes I outlined earlier that Gill (2010) sees Rousseau as emphasizing in the Savoyard Priest section.

put[s] him”⁸⁷ seek to strategically expose Emile to those scenes that will make him morally responsive to his fellows. Rousseau encapsulates this method as follows:

To excite and nourish this nascent sensibility, to guide it or follow it in its natural inclination, what is there to do other than *to offer the young man objects* on which the expansive force of his heart can act – objects which swell the heart, which extend it to other beings, which make it find itself everywhere outside of itself – and carefully to keep away those which contract and concentrate the heart and tighten the spring of the human *I*?⁸⁸

When it comes to the question of determining which objects will extend one’s moral sensibility in this way, on the sensationalist account this is a matter that can only be settled via empirical investigation. One must engage in a meticulous study of the effects of various stimuli on the way people see and feel their relations with others.⁸⁹ If we undertake such observation, Rousseau thinks we will soon discover that some objects extend the heart and attach our affections to others, while others repel us from them. On the strength of his own observations, Rousseau suggests that “all men are affected” and feel some basic impulse of pity when they are in close bodily proximity to the pains of another.⁹⁰ He thus maintains that it is experiences of weakness, vulnerability and “our common miseries which turn our hearts to humanity.”⁹¹ In a statement that would be shared by many other sensationalist and moral sense theorists of the period he writes:

We are attached to our fellows less by the sentiment of their pleasures than by the sentiment of their pains, for we see far better in the latter the identity of our natures with theirs and the guarantees of their attachment to us.⁹²

Having derived from experience this principle that “our common miseries unite us by affection,”⁹³ Rousseau thinks it is only a short step to seeing how it can be employed in deliberate programs of moral cultivation.⁹⁴ In short, the path to learning to see and feel the relations between oneself and others in a distortion-free manner is to study this vulnerability that inescapably characterizes human life. But this is not an academic study, rather Rousseau’s plan for the development of Emile’s moral sensibility revolves around an emphasis on bodily exposure to real instances of such

⁸⁷ Rousseau (2010, 371).

⁸⁸ Ibid., 374–5 (my emphasis).

⁸⁹ Ibid., 379, 364, 370, 410.

⁹⁰ Rousseau writes that experience shows us as a general rule that “all men are affected sooner and more generally by wounds, cries, groans, the apparatus of painful operations, and all that brings objects of suffering to the senses...they are universal, and no one is completely exempt from them” (ibid., 379).

⁹¹ Ibid., 372. That is, experiences of the fact that “men are not naturally Kings, or Lords, or Courtiers, or rich men. All are born naked and poor; all are subject to the miseries of life, to sorrows, ills, needs, and pains of every kind. Finally, all are condemned to death” (ibid., 373).

⁹² Ibid.

⁹³ Ibid.

⁹⁴ As he puts it elsewhere: “On the basis of this principle it is easy to see how all the passions of children and men can be directed to good or bad” (ibid., 364).

suffering. He frequently contends that Emile cannot just be taught this directly but must be intentionally exposed to circumstances (or, at least, must not be intentionally sheltered away from all those inevitable circumstances of human life) in which he might be affected by his fellows' pains. Thus, as Rousseau puts it, "in order to incline a young man to humanity, far from making him admire the brilliant lot of others, *one must show him* the sad sides of that lot, one must make him fear it."⁹⁵ He also admonishes: "Above all, do not go and tell him this coldly like his catechism. *Let him see, let him feel* the human calamities. Unsettle and frighten his imagination with the perils by which every man is constantly surrounded."⁹⁶ And again, "do not get lost in fine reasonings intended to prove to the adolescent that he is a man like others and subject to the same weaknesses. *Make him feel it* or he will never know it"⁹⁷ since "to see it without feeling it is not to know it."⁹⁸ Bodily proximity and affective impact are thus crucial for this education's success.⁹⁹

Of course, given the sort of experiences in question, Rousseau acknowledges that this is a rather controversial program:

Unhappy men, dying ones, sights of pain and misery! What happiness! What enjoyment for a young heart being born to life! His gloomy teacher, who designed so sweet an education for him, treats him as born only to suffer. This is what will be said.¹⁰⁰

Rousseau, however, is confident that his education will be vindicated on the basis of the very different habitual modes of moral perception and feeling it will produce, chief among them being that "he shares the sufferings of his fellows,"¹⁰¹ i.e., that Emile will be able to imaginatively identify and empathize with others; that he will be sensitive, "humane"¹⁰² and "more loving,"¹⁰³ having a "superabundant sensibility he can accord to the suffering of others."¹⁰⁴ Once again then, while outside the scope of this chapter, critically assessing Rousseau's program will hinge on debating the

⁹⁵ Ibid., 375 (my emphasis).

⁹⁶ Ibid., 376 (my emphasis).

⁹⁷ Ibid., 401 (my emphasis).

⁹⁸ Ibid., 373.

⁹⁹ Of course, part of helping Emile come to see and feel "the human calamities" involves working on the various beliefs and conceptions that inform his sentiments since, as Rousseau puts it, "truth of sentiments depends in large measure on correctness of ideas" (ibid., 380). It is thus no more a merely physical program than it is a merely intellectual one. Emile's thoughts and judgments about others play a crucial role in his sensory-affective experience of them, and vice versa. It must be noted, however, that at another point in the text when worrying about potential objections to this method (particularly the risky dangers of overexposure to such scenes), Rousseau steps back from this insistence on bodily proximity by claiming that it is in fact one's *reflection* on any such bodily encounters which is more significant than that experience itself (ibid., 384).

¹⁰⁰ Ibid., 380.

¹⁰¹ Ibid., 382.

¹⁰² Ibid., 377.

¹⁰³ Ibid., 383.

¹⁰⁴ Ibid., 382.

validity of his purported empirical observations regarding the effects of various environments in shaping particular habitual passions, actions etc. – indeed, as noted in Sect. 9.2, this is precisely the sort of developmental inquiry into which he invites his readers through the unique form of his text.

Yet even with his firm conviction regarding the positive outcomes of such an education for Emile’s moral character, Rousseau at the same time offers a highly nuanced account of its complexities. If, after all, the tutor’s role involves managing an entire regime of impression-exposure, intentionally orchestrating everything the child encounters, this will have to be understood as an incredibly precarious and delicate art.¹⁰⁵ And it is perhaps here that Rousseau’s unique contribution emerges in comparison to other key lines of pedagogical thought in the French context. While many in the sensationalist tradition advocated ‘positive education’ programs as a new form of science able to engineer moral citizens with rigorous precision, Rousseau continually emphasizes an infinite variability that would confound such programs:

This is the spirit of the method which must be prescribed. Here examples and details are useless because the almost infinite division of characters begins at this point, and each example I might give would perhaps not be suitable for even one in a hundred thousand. It is also at this age that the skillful master begins to take on the true function of the observer and philosopher who knows *the art of sounding hearts while working to form them*.¹⁰⁶

Rousseau thus argues that managing such an experiential education – “husbanding examples, lessons, and images,”¹⁰⁷ – is a complicated *art*. The tutor must become incredibly attentive to how “every object [he] presents to [Emile]”¹⁰⁸ impresses upon him; how it is schooling him in certain habitual ways of seeing and feeling. Furthermore, as his development proceeds, the tutor must keep making adjustments as required: it is an “art of sounding hearts *while working to form them*.”¹⁰⁹ Rousseau emphasizes the need for vigilant attentiveness regarding these dynamics. For example, if Emile begins to compare himself with others, perhaps feeling that his rank, riches or good health make him invulnerable to the sufferings of those of lower social standing than himself, the tutor, he implies, may well have to up the ante by selecting stronger scenes designed to dispel these illusions of exemption – and, the worse the illusions, the more dramatic the experiential cure.¹¹⁰ However, at the opposite

¹⁰⁵ Much critical work has also focused on exposing the sinister degree of surveillance and manipulation such a controlled program would require.

¹⁰⁶ *Ibid.*, 379 (my emphasis).

¹⁰⁷ *Ibid.*, 384.

¹⁰⁸ *Ibid.*

¹⁰⁹ *Ibid.*, 379 (my emphasis).

¹¹⁰ Here consider for example Rousseau’s commendation of a father’s actions of choosing an “emphatic scene which struck the young man, made an impression on him which was never effaced” (*ibid.*, 385). Describing this general approach Rousseau writes: “To the extent he becomes enlightened, choose ideas which take account of that fact; to the extent his desires catch fire, choose scenes fit to repress them” (*ibid.*, 384–5).

end of the spectrum, now acknowledging the valid concern in the earlier quote above regarding the gloomy subject matter of this education, Rousseau also harbors a profound anxiety about the way overexposure to intense scenes can produce a hypersensitivity that ends only in desensitization and pitiless insensibility.¹¹¹ Once again, this is why understanding that this education is a complicated art is so vital.

Clearly, then, while the method of exposing Emile to the right impressions relies upon certain general principles regarding how different objects as a rule impact upon us, it is ultimately an *art* of cultivation tailored to an individual's unique sensibility. Linking this to the key point made earlier: it is a full program of education involving sensory, cognitive, emotional, and imaginative elements – one which seeks to train Emile in an entire mode of moral perception; an alternative way of being affected and moved to action in his relations with others. That this cultivation concerns a sensibility in which the cognitive and affective are inseparable is further evident in the way Rousseau envisions the final outcome of Emile's sentimental education. The "result of the cultivation of his inclinations" includes both "sublime sentiments" and "judicial clarity" or "accuracy of reason."¹¹² To bring them together we might say Rousseau's goal is for Emile to develop an increasingly judicious sensibility or an increasingly sensitive understanding. The two go hand in hand to produce a sharper and more subtle perception of the moral (and, as he will go on to elaborate, the political)¹¹³ contours of the world outside himself: about the feelings of his fellows, the relations between people, about pragmatic questions concerning what can be done in certain situations to help, how to go about this and so on.¹¹⁴ The aim is that Emile will become increasingly able, as Rousseau sums it up, to *see* "what can do good and what stands in its way."¹¹⁵

It is now clearer the extent to which Rousseau's method is designed to provide an education to rival that offered by society. Given how very good he thinks our social practices are at training us in a distorted, comparative mode of moral perception, those who wish to offer an alternative must likewise devote serious attention to that entire learning process by which the nature of our relation to others is shaped. Furthermore, realizing that this is an education in which there can be no neat separation of cognitive and affective elements, the intentional cultivation of moral sensibility

¹¹¹ The following quote perfectly encapsulates this worry: "The object is not to make your pupil a male nurse or a brother of charity, not to afflict his sight with constant objects of pain and suffering, not to march from sick person to sick person, from hospital to hospital, and from the Grève to the prisons. He must be touched and not hardened by the sight of human miseries. Long struck by the same sights, we no longer feel their impressions. Habit accustoms us to everything. What we see too much, we no longer imagine, and it is only imagination which makes us feel the ills of others. It is thus by dint of seeing death and suffering that Priests and Doctors become pitiless. Therefore, let your pupil know the fate of men and the miseries of his fellows, but do not let him witness them too often" (ibid., 384).

¹¹² Ibid., 410.

¹¹³ On this distinction between morals and politics, see Rousseau (2010, 389).

¹¹⁴ Ibid., 406–7.

¹¹⁵ Ibid., 410.

must likewise capitalize on all these avenues. To merely try a cognitive approach, for example, by teaching Emile some alternative facts, would be to lose out sorely against the education he thinks contemporary Parisian society offers – one which is so skillful at structuring our emotional experience. Hence, we see the motivation for Rousseau’s constant emphasis on moral cultivation being no “esoteric philosophical training”¹¹⁶ but rather a pragmatic program of bodily practices at seeing, feeling and responding beneficently to real others.¹¹⁷ As such, it is hardly likely to be a quick or easy process but rather, as Rousseau nicely puts it, for Emile to study himself in his relations with others is “*the job of his whole life.*”¹¹⁸ The formation of moral sensibility is an ongoing task, an art of cultivation.

In view of this, however, Rousseau also has some rather harsh things to say about those unwilling to take the time, care and effort to engage in this art. On his account, tutors who neglect it are culpable for the distorted moral sensibilities of their pupils. Likewise, Rousseau also seems to imply that for an adult any claims, for instance, of a supposed inability to feel empathy – to use one of his own examples, that seen in the hardness of the rich towards the poor¹¹⁹ – are likewise blameworthy. They result from a similar neglect of such cultivation, a failure to put the time and effort into becoming “master of directing [one’s] imagination toward this or that object or of giving it this or that habit.”¹²⁰ Thus, while Rousseau is certainly happy to acknowledge that a person’s capacity to be affected by the impressions of another’s pain “have their modifications and their degrees which depend on the particular character of each individual and his previous habits,”¹²¹ nevertheless, since these habits are themselves susceptible to re-habitation (though the longer one leaves it, the harder this will get), he also sees fit to judge people blameworthy for their current empathetic ineptitude.

Even if then we ultimately end up parting company with Rousseau over his portrait of the end goal of such education (i.e., his specific criteria regarding what counts as naturalness in society – a topic which has largely been left to one side in this chapter), there is much in his extensive engagement with the processes of this education that might still remain instructive. In particular, Rousseau’s investigations into the type of problem at stake in, what he deems to be, his contemporaries’ distorted moral sensibilities contains insights applicable to any analysis concerned with cases of deeply entrenched, potentially distorted ways of seeing and feeling or, as Kant formulated it in his own case, with the sorts of “blinding prejudice” that can come to shape one’s entire manner of relating to others. This chapter has highlighted

¹¹⁶ Gaukroger (2010, 419).

¹¹⁷ For Rousseau’s emphasis on a program of active beneficence and learning to be good by doing good, see Rousseau (2010, 406–7).

¹¹⁸ *Ibid.*, 364 (my emphasis).

¹¹⁹ *Ibid.*, 376.

¹²⁰ *Ibid.*, 371. We can perhaps see an analogue of this in Rousseau’s odd little discussion of the physical agility of children, where he writes: “It seems to me, that the supposed ineptitude of children at our exercises is imaginary and that, if they are not seen to succeed at some, it is because they have never been given practice in them” (Rousseau 2010, 290).

¹²¹ *Ibid.*, 379.

Rousseau's attentiveness to the breadth and depth of such distortions, his investigations into their source in one's social environment and his own educational experiment in the *Emile* to create the sentimental conditions that would enable another – in his view, more truthful – “way of being, of seeing, and of feeling.”¹²² Through the course of his analyses Rousseau emerges as an astute observer of the malleability and stubbornness of psychological formations at different developmental stages, of the complex interplay of the cognitive and the affective in this development and of the opportunities further experience and educational intervention might or might not be able to offer for re-forming what has already taken decisive shape. In these respects then Rousseau's rhapsodic fidelity to the task of cultivating the man of nature can be seen to have far less in common with utopian dreams of withdrawing to “the depths of the woods”¹²³ than it does with sophisticated philosophical inquiry into questions of moral cultivation.

References

- Cooper, Laurence D. 1999. *Rousseau, nature, and the problem of the good life*. University Park: The Pennsylvania State University Press.
- Dent, N.J.H. 1988. *Rousseau: An introduction to his psychological, social and political theory*. Oxford: Blackwell.
- Dent, Nicholas. 2005. *Rousseau*. New York: Routledge.
- Gaukroger, Stephen. 2010. *The collapse of mechanism and the rise of sensibility: Science and the shaping of modernity, 1680–1760*. Oxford: Oxford University Press.
- Gill, Natasha. 2010. *Educational philosophy in the French enlightenment: From nature to second nature*. Farnham/Burlington: Ashgate Publishing.
- Grimsley, Ronald. 1983. *Jean-Jacques Rousseau*. Brighton: Harvester Press.
- Johnston, Steven. 1999. *Encountering tragedy: Rousseau and the project of democratic order*. New York: Cornell University Press.
- Lloyd, Genevieve. 1983. Rousseau on reason, nature and women. *Metaphilosophy* 14(3–4): 308–326.
- Marks, Jonathan. 2005. *Perfection and disharmony in the thought of Jean-Jacques Rousseau*. Cambridge: Cambridge University Press.
- O'Neal, John C. 1985. *Seeing and observing: Rousseau's rhetoric of perception*. Saratoga: ANMA Libri.
- Rorty, Amélie Oksenberg. 1998. Rousseau's educational experiments. In *Philosophers on education: Historical perspectives*, ed. Amélie Oksenberg Rorty, 237–253. London: Routledge.
- Rousseau, Jean-Jacques. 1990. *Rousseau, judge of Jean-Jacques, dialogues*. Trans. Judith R. Bush, Christopher Kelly, and Roger D. Masters, ed. Roger D. Masters, and Christopher Kelly. Hanover/London: University Press of New England.
- Rousseau, Jean-Jacques. 2010. *Emile or on education (includes Emile and Sophie, or the solitaires)*. ed. and Trans. Christopher Kelly, and Allan Bloom. Lebanon: University Press of New England.
- Scott, John T. 2006. *Jean-Jacques Rousseau: Human nature and history*. Abingdon/New York: Routledge.
- Vila, Anne C. 1998. *Enlightenment and pathology: Sensibility in the literature and medicine of eighteenth-century France*. Baltimore: Johns Hopkins University Press.

¹²² *Ibid.*, 401.

¹²³ Rousseau (2010, 412).

Chapter 10

The Piacular, or on Seeing Oneself as a Moral Cause in Adam Smith

Eric Schliesser

10.1 Introduction and Theses

Consider the following passage:

A man of humanity, who accidentally, and without the smallest degree of blamable negligence, has been the cause of the death of another man, feels himself piacular, though not guilty. During his whole life he considers this accident as one of the greatest misfortunes that could have befallen him. If the family of the slain is poor, and he himself in tolerable circumstances, he immediately takes them under his protection, and, without any other merit, thinks them entitled to every degree of favour and kindness. If they are in better circumstances, he endeavours by every submission, by every expression of sorrow, by rendering them every good office which he can devise or they accept of, to atone for what has happened, and to propitiate, as much as possible, their, perhaps natural, though no doubt most unjust resentment, for the great, though involuntary, offence which he has given them. (*The Theory of Moral Sentiments*, 2.3.3.4, 107)¹

In this paper, I explore the significance of that peculiar concept, the so-called piacular, in Adam Smith's *The Theory of Moral Sentiments* (hereafter TMS). The three main paragraphs of TMS that I discuss below (2.3.3.4, 2.3.3.5, 7.4.30) were all added to the final, sixth edition. Smith describes the concept first in the context of his treatment of what we would call "moral luck" and then returns to it in what became part VII of TMS.² In brief, the piacular is the feeling that arises when we have been an involuntary cause of another's harm. It is a feeling of shame that is akin – but not identical – to what is commonly called "agent-regret."³ In its discomfort

¹ I quote from the Glasgow Edition of all of Smith's works by paragraph and page-numbers.

² For the ongoing debate over Smith's views on moral luck, see Russell (1999), Garrett (2004), Flanders (2006), and Hankins (ms). Of these Flanders and Hankins discuss the piacular.

³ Williams (1976).

E. Schliesser (✉)
Department of Philosophy and Moral Sciences,
University of Ghent, 9000 Ghent, Belgium
e-mail: Nescio2@yahoo.com

the piacular motivates compensatory behavior of atonement that is governed by highly specific norms. In Williams' treatment, agent-regret also motivates atonement. In what follows, while unpacking Smith's understanding of the piacular, I occasionally remark on the ways in which it is similar to agent-regret. I do so in order to highlight the crucial normative differences between the Smithian piacular and agent-regret.

I argue, first, that according to Smith it is part of our humanity that we ought to see ourselves in part as causes in the (great) causal chain of life. This is a plausible interpretation of Smith's view in light of (i) his treatment of the way in which the sympathetic process that underwrites moral judgment is, in part, a judgment of the proportionality between causes and effects and (ii) his claim that our habitual causal environment is constitutive of our sanity and rationality. Second, I explain the norms that according to Smith govern the atonement of the piacular. Somewhat surprisingly, these norms are irrevocably tainted by superstition. In Smith's account this superstitious element should not be eradicated, but embraced as part of our shared humanity.

My treatment is primarily exegetical – I aim to fit the piacular into Smith's other reflections on nature and normativity. I also do not engage with the debate concerning the extent to which Smith's views get "moral luck" right. My primary focus is on how we should think about the piacular and the norms governing it. Moreover, I cannot appeal to the natural sympathy of those already convinced by sentimentalist moral theorizing, of which Smith is often treated as a founding father,⁴ because here I focus on idiosyncratic aspects of his moral phenomenology not shared by recent philosophy. Having said that, his views capture aspects of our moral lives that we have a tendency to wish or explain away.

In what follows I proceed as follows: in Sects. 10.2 and 10.3 I, first, explain how knowledge of causation plays a constitutive role in Smithian moral judgment; second, I explain how causation and rationality hang together in Smith's system. I then turn to a detailed account of the piacular in Sects. 10.4 and 10.5. I focus on explaining why, according to Smith, it is an instance of shame as well as why, in explicating this shame, Smith is so willing to flirt with superstition (Sect. 10.6). In the conclusion I explore the significance of the argument presented here via an examination of some objections that may be raised against it.

10.2 Sympathy and Knowledge of Causal Relations⁵

Consider the following passage in TMS:

There are some passions of which the expressions excite no sort of sympathy, but before we are acquainted with what gave occasion to them, serve rather to disgust and provoke us

⁴ See, e.g., Gibbard (1990, chapters 9 and 15).

⁵ The material in this section is excerpted with minor revisions from a forthcoming paper, Schliesser (Forthcoming).

against them ... The general idea of good or bad fortune, therefore, creates some concern for the person who has met with it, but the general idea of provocation excites no sympathy with the anger of the man who has received it. Nature, it seems, teaches us to be more averse to enter into this passion, and, till informed of its cause, to be disposed rather to take part against it. Even our sympathy with the grief or joy of another, before we are informed of the cause of either, is always extremely imperfect. (TMS 1.1.1.7-9, 11)

In context, Smith's claim is that what we might label 'instinctual sympathy' is limited. In cases where we lack knowledge of the causal circumstances responsible for an agent's passions, the sympathetic process will always lead to what Smith calls "imperfect sympathy." It seems to follow from his terminology that there exists some (non-instinctual) sequence that leads to perfect, or at least much less imperfect, sympathy. The sympathetic process can be broken down in the following (largely sequential) steps: following (T0) (i) an intensely/passionately felt moral situation, which is experienced or observed empirically, consists in the following: (ii) by way of the imagination spectators and moral agents place themselves in each other's situations, including (iii) knowledge of the (moral) causes that gave rise to the moral situation. This involves (iv) a sympathetic mutual modulation (informed, perhaps, by observations about how the other is reacting), which, in turn, produces (v) a conceived, reflected passion within each participant in the sympathetic process. This then (vi) alters the intensity of the feelings of the participants in the process. After several rounds of this, perhaps, it produces (vii) fellow feeling (sympathy) between the spectator and the other persons principally concerned.

It is crucial for my approach to Smith that (iii) is not ad hoc but in fact reflects significant currents in Smith's thinking about moral evaluation and moral agency. Consider a standard summary that Smith provides about the natures of propriety and impropriety, on the one hand, and merit or demerit, on the other hand:

It has already been observed [TMS 1.1.3.5, 18; ES], that the sentiment or affection of the heart, from which any action proceeds, and upon which its whole virtue or vice depends, may be considered under two different aspects, or in two different relations: first, in relation to the cause or object which excites it; and, secondly, in relation to the end which it proposes, or to the effect which it tends to produce: that upon the suitability or unsuitability, upon the proportion or disproportion, which the affection seems to bear to the cause or object which excites it, depends the propriety or impropriety, the decency or ungracefulness of the consequent action; and that upon the beneficial or hurtful effects which the affection proposes or tends to produce, depends the merit or demerit, the good or ill desert of the action to which it gives occasion. (TMS 2.1. Intro.2, 67)

There is a lot going on here and I am not going to provide even the semblance of a full treatment of the Smithian concept of propriety (or merit). All I want to suggest is that causal relations are constitutive of the nature of both Smithian propriety and merit. That is, the two central Smithian moral judgments can be characterized schematically according to the following temporal sequence: (u) an exciting cause, which produces (v) a sentiment of heart, which leads to (w) an action and (x) its foreseeable effects, and, of course, (y) the actual effects produced by (w). Now, first, judgments of propriety and impropriety, which are principally concerned with situations, are judgments regarding the proportion among (u)-(v)-(w)-(x). Meanwhile, second, judgments of merit and demerit, which are fundamentally judgments of character,

focus on the proportion among (v)-(w)-(y). These two sequences are fundamentally causal in nature (i.e., “excites,” the “effects” it tends “to produce,” etc.).

Of course, the previous paragraph is a gross simplification and ignores considerable complexity in Smith’s treatment of propriety (etc.). However, it is crucial to my argument that Smith introduces causal language in describing the content of our moral judgments. In particular, he suggests that when we make a moral judgment we do so only after mentally inspecting, as it were, the proportionality of the relata that enter into a cause-effect relation.

It might be thought that Smith’s terminology here is evidence of his Humean debts. Indeed, it echoes the manner in which Hume treats the natural relation of cause and effect at *Treatise* 1.1.4. Initially then there can be no doubt that Smith is deploying a Humean framework regarding the nature of causation here – one in which causes are regular successions of a certain sort. However, upon reflection, Smith also subtly diverges from Hume in that Hume accepts the following position: “An effect always holds proportion with its cause” (“Of Interest,” Hume 1987: 297). Elsewhere, I have dubbed this “Hume’s ninth rule” (because it follows from the conjunction of Hume’s fourth and seventh (out of eight) “rules by which to judge of causes and effects” (*Treatise* 1.3.15)).⁶ So, Smith’s position is that, when Hume’s ninth rule obtains, we are inclined to make judgments of propriety and merit. However, in Smith’s approach to our moral life, the ninth rule regularly need not hold. After all, it is not as if judgments of impropriety need to be rare according to Smith. Rather, he contends that moral causes and effects can be monstrously out of proportion – a claim that is, in fact, much harder to incorporate into Hume’s framework.

I now turn to Smith’s views on causation in order to explain the tight link between causation and rationality in his thought.

10.3 Causation and Rationality

As is well known, Smith had most of his papers destroyed before his death. But throughout his life he wished that a number of surviving manuscripts would be published in case of his premature death. These were eventually published posthumously in (the 1795) *Essays on Philosophical Subjects* (EPS) The lengthiest of these pieces is the “History of Astronomy,” (hereafter Astronomy) which continues to be studied by those interested in Smith’s philosophy of science and, because of its uncanny anticipations of the ideas of Thomas Kuhn, those working on the pre-history of historical approaches to philosophy of science.⁷ Consider this passage from this text:

It is evident that the mind takes pleasure in observing the resemblances that are discoverable betwixt different objects. It is by means of such observations that it endeavours to

⁶ I quote from Hume (2004) by paragraph number.

⁷ For an introduction see Berry (2006); see also Schliesser (2005).

arrange and methodise all its ideas, and to reduce them into proper classes and assortments ... When two objects, however unlike, have often been observed to follow each other, and have constantly presented themselves to the senses in that order, they come to be so connected together in the fancy, that the idea of the one seems, of its own accord, to call up and introduce that of the other. If the objects are still observed to succeed each other as before, this connection, or, as it has been called, this association of their ideas, becomes stricter and stricter, and the habit of the imagination to pass from the conception of the one to that of the other, grows more and more rivetted and confirmed. As its ideas move more rapidly than external objects, it is continually running before them, and therefore anticipates, before it happens, every event which falls out according to this ordinary course of things. (Astronomy, 2.2-7, 38–41)

Smithian causation is founded on a pleasing psychological disposition that activates the mind to notice and, perhaps, even search out resemblances between objects. This pleasing activity of classifying ideas leads into what we may call a natural taxonomy including abstract categories. (In a piece, “Considerations concerning the First Formation of Languages,” that Smith attached to the third edition of TMS, Smith had, while engaging Rousseau on the question of the origin of languages, offered a how-possible articulation of the growth of this disposition in which mind and language co-evolve toward increasing metaphysical complexity through historical time.)⁸ In the most Humean part of Smith’s treatment above, the observation of the ordered, constant conjunction of such habitual resemblance leads to the association of ideas and the effortless – to introduce Humean terminology – movement of the natural relation, which occurs (as Smith says) in the imagination.

Now, in my description above I have used “causation” (even though Smith does not),⁹ because (i) Smith’s calls attention to the Humean provenance of the view and (ii) the relation described fits Hume’s *Treatise* (projective-) regularity account of causation, which emphasizes how the mind acquires the feeling of necessity due to the constant conjunction of two objects and the priority of one (the cause) over the other (the effect). Having said that, Smith does not mention the contiguity requirement that is present in the *Treatise* (and dropped in Hume’s first *Enquiry*). Smith’s final move in the passage above is also not to be found in Hume (although itself not un-Humean). For Smith, once associated, our ideas move more rapidly than the (external) objects that originated them.¹⁰ In adults this process creates a never-ending stream of mental anticipations of the world. When the world deviates from these anticipations this generates a painful emotion of wonder and the desire to alleviate that feeling, which, in turn, leads (and this is crucial for the larger aim of the “History

⁸ In the Glasgow Edition it can be found in Smith *Lectures on Rhetoric and Belles Lettres*. For discussion, see Berry (1974), Otteson (2002), Levy (1997), and Schliesser (2011).

⁹ In particular, here I am agnostic about to what degree “constant conjunction of such habitual resemblance” always tracks or detects metaphysical causation in nature according to Smith. To settle this question, however, one must decide to what degree Smith is a so-called skeptical realist (as Hanley 2010 thinks) or a (culturally sensitive) modest realist (as Montes 2004 and Schliesser 2005 think). It is possible, as Hanley suggests, that Smith is a skeptical realist only in the moral sciences and not in physical sciences. I thank Andrew Corsa for this discussion.

¹⁰ Andy Clark’s (2012, in press) Bayesian updating of the model shares surprising features with Smith’s approach.

of Astronomy”) in some circumstances to the desire to inquire into the world. The more regular response to such events, however, is

the origin of Polytheism, and of that vulgar superstition which ascribes all the irregular events of nature to the favour or displeasure of intelligent, though invisible beings, to gods, daemons, witches, genii, fairies. For it may be observed, that in all Polytheistic religions, among savages, as well as in the early ages of Heathen antiquity, it is the irregular events of nature only that are ascribed to the agency and power of their gods. (Astronomy, 3.2, 49)¹¹

Let me now turn to a Smithian thought-experiment, which radicalizes one of Hume’s¹² and that gets us to the nub of this section:

Could we conceive a person of the soundest judgment, who had grown up to maturity, and whose imagination had acquired those habits, and that mold, which the constitution of things in this world necessarily impress upon it, to be all at once transported alive to some other planet, where nature was governed by laws quite different from those which take place here; as he would be continually obliged to attend to events, which must to him appear in the highest degree jarring, irregular, and discordant, he would soon feel the same confusion and giddiness begin to come upon him, which would at last end in the same manner, in lunacy and distraction. Neither, to produce this effect, is it necessary that the objects should be either great or interesting, or even uncommon, in themselves. It is sufficient that they follow one another in an uncommon order. (Astronomy, 2.10, 43)

There are two aspects about this thought-experiment that are significant here: first, sound judgment seems to consist in having the right kind of habits, that is, ones that match the world’s natural order. It is unclear to what degree we have any control over these habits.¹³ At any rate, for Smith, in a properly functioning person there is a reciprocal relationship between the habituated mental anticipations and sound judgment. Second, on Smith’s view (one version, perhaps, of) lunacy just is having mental anticipations that are systematically out of kilter with the order in which objects appear. That is, when our firmest, most habitual, natural relations of causation closely track our (natural or) common environment this is constitutive of (mental) sanity and rationality. In fact, in Smith’s treatment of the impact of custom and education among so-called civilized societies (that is, in law-governed places), he describes immoral behavior in language that echoes this thought-experiment: one is “transported to do any thing contrary to justice or humanity” (TMS 5.2.10, 207). It is as if he thinks of immoral behavior in terms of one’s being an alien in one’s environment. It may seem

¹¹ As an aside, from the context it is unclear what, according to Smith, will count as refined superstition.

¹² “It may not be amiss to observe on this occasion, that the influence of general rules and maxims on the passions very much contributes to facilitate the effects of all the principles, which we shall explain in the progress of this treatise. For it is evident, that if a person full-grown, and of the same nature with ourselves, were on a sudden-transported into our world, he would be very much embarrassed with every object, and would not readily find what degree of love or hatred, pride or humility, or any other passion he ought to attribute to it. The passions are often varied by very inconsiderable principles; and these do not always play with a perfect regularity, especially on the first trial. But as custom and practice have brought to light all these principles, and have settled the just value of every thing.” (Hume *Treatise*, 2.1.6.9)

¹³ Hume’s “A wise man, therefore, proportions his belief to the evidence,” (EHU 10.1.4) seems to have to do more work to get it right.

that my use of “rationality” here is unwarranted, however I take it that the “person of soundest judgment” *is* the rational person. Of course, as Smith makes clear, sometimes judgments of proportionality require distance: “I can form a just comparison . . . in no other way, than by transporting myself, at least in fancy, to a different station, from whence I can survey both at nearly equal distances, and thereby form some judgment of their real proportions.” (TMS 3.3.2, 135)

If we now connect these two aspects to the earlier observation that when we make a moral judgment we do so only, after mentally inspecting, as it were, the proportionality of the relata that enter into the cause-effect relation, we can infer that on Smith’s view, when we are functioning properly we have been attuned in a stable, well-ordered environment to make such proper judgments.¹⁴

The two previous sections have claimed that in Smith’s thought there is a surprisingly tight link among nature’s regularity, our rationality as well as sanity, and moral judgment. Even the sympathetic process itself is not only a causal process but also relies on our ability to discern causal relations. Thus, in Smith’s thinking about morality the very fact that we are fundamentally a part of the chain of causes plays a non-trivial role. I now turn to a more focused discussion of the piacular.

10.4 We (Ought to) See Ourselves as Causes!

Smith writes:

A man of humanity, who accidentally, and without the smallest degree of blamable negligence, has been the cause of the death of another man, feels himself piacular, though not guilty. During his whole life he considers this accident as one of the greatest misfortunes that could have befallen him. If the family of the slain is poor, and he himself in tolerable circumstances, he immediately takes them under his protection, and, without any other merit, thinks them entitled to every degree of favour and kindness. If they are in better circumstances, he endeavours by every submission, by every expression of sorrow, by rendering them every good office which he can devise or they accept of, to atone for what has happened, and to propitiate, as much as possible, their, perhaps natural, though no doubt most unjust resentment, for the great, though involuntary, offence which he has given them. (TMS 2.3.3.4, 107)

As has been argued thus far, if we are part of a voluntary cause-effect sequence, then the categories of propriety/guilt and merit/demerit are appropriate. If, however, we find ourselves part of an involuntary cause-effect sequence then the category of the piacular can be appropriate¹⁵ – after all, it is predicated of “the man of humanity.” In some other contexts “the man of humanity” is treated as akin to the right kind of

¹⁴ Obviously, here I am gesturing at a much larger argument concerning moral education in Smith.

¹⁵ Given that when one is an involuntary cause one’s motive does not matter, there is only one category here. Of course, one may be responsible for activity further back in the causal chain, but that is beside the point here.

impartial spectator (see especially, TMS 2.2.3.11, 90 & 3.3.4, 137 and 3.3.26, 147), including non-trivial criticism of Hume's notorious treatment of scratching of one's finger at *Treatise* 2.3.3.¹⁶ Of course, this is not to claim that all men of humanity are praiseworthy, according to Smith, "we so frequently find in the world men of great humanity who have little self-command, but who are indolent and irresolute, and easily disheartened, either by difficulty or danger, from the most honourable pursuits" (TMS 3.3.37, 153).¹⁷ In the case in which the man of humanity is an involuntary cause, his response is not faulted by Smith in any fashion.¹⁸ This can be seen in the above example in the way Smith explicitly treats the resentful response of the victim's family as "most unjust," even though their feeling is quite "natural." By contrast, if the man of humanity had been a voluntary cause of the victim's death, their feeling of resentment would have been just.¹⁹ One of the useful aspects of the piacular feeling then is to "learn to dread that animal resentment" (2.3.3.4, 106).

We are also led to think that the feelings of the man of humanity are perhaps not so natural – at least in the sense that most of us (who are probably not regularly men or women of humanity) will mistakenly try to use our lack of culpability or blameworthiness as a way to avoid the feeling of the piacular (I return to this below). Even so, as earlier commentators have emphasized, echoing Bernard Williams, if we lacked any feeling whatsoever about even the involuntary harm caused to others by us, something would be remiss.²⁰ That is to say, we are causes and it is part of our humanity that we ought to understand ourselves as such.²¹ As an aside, in its acknowledgment of the significance of outside causes to our well-being, this is a most un-Stoic move in Smith's thought.

Smith appropriated an old concept for this feeling in labeling it the "piacular". The term is derived from the Latin "*piāculum*", which means propitiatory sacrifice, and from "*piāre*", which means to appease. In Smith, this presumably unpleasant feeling we ought to have when we are the involuntary cause of harm to others, leads, as its etymology suggests, to compensatory behavior (as it also does in Williams' case of "agent-regret"). The norms that govern this appeasement have not been much explored hitherto. In what follows, I first describe these norms and then turn to consider why they help explain, in part, Smith's need for the use of this concept and not, say, guilt or regret.

¹⁶ See Fleischacker (1999: 42). For the significance of humanity as a moral category in Smith (and Hume), see Taylor (2006), Debes (2007), and Hanley (2011).

¹⁷ Their possible lack of self-command suggests that not all men of humanity have perfect virtue. See also TMS 3.3.35, 151, and 1.1.5.1, 23; I thank Ryan Hanley and Andrew Corsa for these references and discussion.

¹⁸ See Flanders (2006: 208–9).

¹⁹ See Pack and Schliesser (2006).

²⁰ As Hankins (ms) notes there is a further debate in the literature regarding whether this feeling is appropriately moral.

²¹ I believe that for Smith this is a post-Enlightenment thought. See Schliesser (2006) and the last section of this paper for further clarification.

10.5 Norms of Appeasement

When Smith first introduces the piacular, he does so with a dramatic case – one is the involuntary cause of the death of another. But as becomes clear in the very next paragraph, many other cases of causing involuntary harm to others can induce the piacular feeling. As Smith writes:

The distress which an innocent person feels, who, by some accident, has been led to do something which, if it had been done with knowledge and design, would have justly exposed him to the deepest reproach, has given occasion to some of the finest and most interesting scenes both of the ancient and of the modern drama. It is this fallacious sense of guilt, if I may call it so, which constitutes the whole distress of Oedipus and Jocasta upon the Greek, of Monimia and Isabella upon the English, theatre. They are all of them in the highest degree piacular, though not one of them is in the smallest degree guilty. (TMS 2.3.3.5, 107)

Here Smith describes three different piacular plays. All four named protagonists in these plays violate pre-established marriage rules and all repay their debt by suicidal death or, in Oedipus' case, with blindness and the giving up of political power. However, the involuntary harm caused is not always death. For example, Jocasta's crime is unknowing incest, while Isabella's is unknowing adultery. Even so, all four are "in the highest degree piacular." This suggests, first, that the piacular comes in degrees. But, second, that perhaps above a certain threshold of harm (for instance, the violation of a major social institution), the piacular becomes maximal.

Furthermore, TMS 2.3.3.4 makes clear that it is not just the harm caused that matters but also the relative social status between the cause of harm and the victim of harm. This suggests that the discharge or the atonement of the piacular feeling is sensitive to social context. However, in the dramatic cases described in 2.3.3.5, the discharge involves death or (sacrificial) suicide by the female characters. From this I infer that, on the one hand, when the harm caused is primarily (or solely) to others the 'price' of atonement is governed, in part, by the relative social status of the victim and the involuntary cause. On the other hand, when the harm caused is a violation against a fundamental institution of civil society (marriage and family laws, etc.) then the 'price' required is the ultimate sacrifice (at least on the stage).

Smith returns to Piacular in part VII of TMS, and his treatment here sheds some further light on the norms governing atonement:

It is not always so with the man, who, from false information, from inadvertency, from precipitancy and rashness, has involuntarily deceived. Though it should be in a matter of little consequence, in telling a piece of common news, for example, if he is a real lover of truth, he is ashamed of his own carelessness, and never fails to embrace the first opportunity of making the fullest acknowledgments. If it is in a matter of some consequence, his contribution is still greater; and if any unlucky or fatal consequence has followed from his misinformation, he can scarce ever forgive himself. Though not guilty, he feels himself to be in the highest degree, what the ancients called, piacular, and is anxious and eager to make every sort of atonement in his power. Such a person might frequently be disposed to lay his case before the casuists, who have in general been very favourable to him, and though they have sometimes justly condemned him for rashness, they have universally acquitted him of the ignominy of falsehood. (TMS 7.4.30, 338–9)

This paragraph is offered in the midst of Smith's critical treatment of casuistry. I offer seven observations: first, this paragraph confirms that the piacular comes on a continuum that is governed by the harm caused. Smith implies that we can make relatively clear judgments about differences in the harm caused. If the harm is fatal to others, one is piacular in the highest degree. This suggests that the norms governing the price of the atonement are determined by the kind of harm (to norms of society, others, self, etc.) and the extent of that harm (how fatal, etc.) – both of which take precedence over the relative social status of victim and involuntary cause.

Second, Smith does not treat the feeling of the piacular as a mistake in some way. Rather, here the “real lover of truth” embraces the feeling. It is hard to imagine that Smith would use this description for somebody fundamentally wrong in his or her reactive attitudes.²² As such, it seems that both the virtuous (“man of humanity”) and the wise (“lover of truth”) can have this feeling. In fact, as I remarked above, I suspect that Smith thinks it is in some sense unnatural to feel piacular. When we are neither virtuous nor wise, we are more likely to seek out the clerical experts (casuists), who – on payment – might sometimes be able to provide us with plausible reasons to evade feeling what we ought to feel.²³ That is to say, ordinarily we sense that we ought to feel something – that there is an apt response – to situations producing harm to others in which we have been an unwilling cause. And yet, unless virtuous and wise, we also tend to go out of our way to avoid feeling it. Given this, part of the indictment against casuistry concerns the way in which it commercially facilitates moral-psychological escapism.²⁴ In contrast to many of his modern admirers Smith does not think that all free markets are always just.

Third, while here Smith does not further elaborate on the significance of relative social status in governing the norms of atonement, he does re-emphasize that in the worst cases atonement will never fully succeed in discharging the piacular feeling. In such cases one “can scarce ever forgive” oneself and, recall, during one's “whole life” one considers “the accident as one of the greatest misfortunes that could have befallen” oneself. This thought extends beyond the pagan origins of the piacular, where “proper atonement” can end the “vengeance of that powerful and invisible being.” (TMS 2.3.3.4, 107)

²² It is the case, however, that if the lover of truth unintentionally misleads on a more regular basis, his fitness to lead can be called into question (TMS 7.4.27, 337). Not all philosophers are suited to be rulers.

²³ While in many contexts Smith acknowledges that casuistry and jurisprudence are related and also admits that some philosophers he admires (Hutcheson and Cicero) were also casuists, fundamentally his judgment of casuistry is on the whole very negative: “Books of casuistry, therefore, are generally as useless as they are commonly tiresome.” (TMS 7.4.33, 339) Perhaps there are a few useful books of casuistry, but I have been unable to find a single mention in Smith.

²⁴ To be clear, Smith's text seems indeterminate between two possible positions: (a) we ordinarily feel some unpleasant inward feeling of shame and seek out casuists to discharge this feeling in order to avoid feeling polluted and (b) we ordinarily know that we ought to feel some unpleasant inward feeling of shame and seek out casuists to prevent feeling even this. I thank Anik Waldow for discussion on this issue.

Fourth, we learn from this passage why Smith is anxious to distinguish the piacular feeling from guilt, and why earlier he had called it a “fallacious sense of guilt.”²⁵ The reason for this is that the piacular is a species of shame. This sets Smith clearly apart from a Williams-inspired reading in terms of agent-regret (which does not invoke shame in Williams’ treatment). Now, while Smith never defines shame, for him it is primarily a consequence of the inward judgment of the imagined impartial spectator – even in an otherwise innocent person (see, for example, his treatment of the case of Calas at TMS 3.2.11, 120). In particular, in Smith, shame is connected to ideas of being permanently stained or polluted.²⁶ Given that the virtuous (“man of humanity”) and the wise (“lover of truth”) can have this feeling, I see no reason to doubt that the impartial spectator can endorse piacular judgments.

Fifth, the third and fourth reasons help us better understand how some of Smith’s tacit reflections on the piacular can be connected to the material covered in Sects. 10.2 and 10.3 of this chapter. If it is constitutive of human being to be, in part, a cause and, furthermore, if all causes are governed by at least the feeling of necessity, then it may well be the case that some lucky ‘causes’ will simply get to glide through life. Meanwhile, other ‘causes’, who may be largely indistinguishable from their fellow ‘causes’, will be marked just in virtue of the harms that follow from their existence. While much of TMS and the account of the “irregularity of sentiments” in particular is quite compatible with and, at first blush, even seems explicitly designed to elicit the approval of providential arguments (for example, in context, Smith frequently talks of Nature’s intentions, the Author of nature’s plan, the “wisdom and goodness of God,” etc.), in this context he is explicit that it is

²⁵ Flanders and Hankins both explore this issue. My explanation is different from theirs. For an important treatment of Smithian guilt, see Brissenden (1969). I thank Ryan Hanley for calling my attention to Brissenden.

²⁶ See, for example, this paragraph which is highly relevant to Smith’s treatment of the piacular above: “Our imagination therefore attaches the idea of shame to all violations of faith, in every circumstance and in every situation. They resemble, in this respect, the violations of chastity in the fair sex, a virtue of which, for the like reasons, we are excessively jealous; and our sentiments are not more delicate with regard to the one, than with regard to the other. Breach of chastity dishonours irretrievably. No circumstances, no solicitation can excuse it; no sorrow, no repentance atone for it. We are so nice in this respect that even a rape dishonours, and the innocence of the mind cannot, in our imagination, wash out the pollution of the body. It is the same case with the violation of faith, when it has been solemnly pledged, even to the most worthless of mankind. Fidelity is so necessary a virtue, that we apprehend it in general to be due even to those to whom nothing else is due, and whom we think it lawful to kill and destroy. It is to no purpose that the person who has been guilty of the breach of it, urges that he promised in order to save his life, and that he broke his promise because it was inconsistent with some other respectable duty to keep it. These circumstances may alleviate, but cannot entirely wipe out his dishonour. He appears to have been guilty of an action with which, in the imaginations of men, some degree of shame is inseparably connected. He has broke a promise which he had solemnly averred he would maintain; and his character, if not irretrievably stained and polluted, has at least a ridicule affixed to it, which it will be very difficult entirely to efface; and no man, I imagine, who had gone through an adventure of this kind would be fond of telling the story.” (TMS 7.4.13, 332–3)

“Fortune, which governs the world.” (TMS 2.3.1, 104).²⁷ The piacular points toward an acknowledgment of the Epicurean system, which was then commonly associated with the rule of chance or Spinozist necessitarianism – both of which are compatible with the rule of Fortune.²⁸

Sixth, cases of the piacular are in Smith’s sense marked by a disproportion between cause and effect and, the worse the case of the piacular, the more this is so. They thus represent a very real danger to the mental health of well-formed minds. In this sense, the most piacular events are very much like the cases Smith describes of sudden, unexpected events:

The passion is then poured in all at once upon the heart, which is thrown, if it is a strong passion, into the most violent and convulsive emotions, such as sometimes cause immediate death; sometimes, by the suddenness of the ecstasy, so entirely disjoint the whole frame of the imagination, that it never after returns to its former tone and composure, but falls either into a frenzy or habitual lunacy; and such as almost always occasion a momentary loss of reason, or of that attention to other things which our situation or our duty requires. (Astronomy, 1.2, 34–5)

So, seventh, all of this tells us that piacular atonement serves three, related purposes: (i) the appeasement of the resentment of the victim (and her family); (ii) the potentially futile attempt at discharging the unpleasant inward feeling of shame; (iii) acknowledging that one is, in some sense, a marked or polluted cause.

It would be tempting to conclude here. However, we have not yet explained why in reflecting on this atonement Smith introduces the language of sacrifice and why he claims an ancient pedigree for the piacular. To answer these questions I turn to the general mechanism of which the piacular is an instance.

10.6 The Language of Superstition

In the context of Smith’s second treatment of the piacular in TMS 7, Smith describes the general psychological mechanism of which it is an instance:

The consciousness, or even the suspicion of having done wrong, is a load upon every mind, and is accompanied with anxiety and terror in all those who are not hardened by long habits of iniquity. Men, in this, as in all other distresses, are naturally eager to disburthen themselves of the oppression which they feel upon their thoughts, by unbosoming the agony of their mind to some person whose secrecy and discretion they can confide in. The shame, which they suffer from this acknowledgment, is fully compensated by that alleviation of their uneasiness which the sympathy of their confident seldom fails to occasion. It relieves

²⁷ While throughout TMS Smith mentions fortune in passing, the most significant other mention is in his famous criticism of the slave-trade: “Fortune never exerted more cruelly her empire over mankind, than when she subjected those nations of heroes to the refuse of the jails of Europe, to wretches who possess the virtues neither of the countries which they come from, nor of those which they go to, and whose levity, brutality, and baseness, so justly expose them to the contempt of the vanquished.” (TMS 5.2.9, 206–7)

²⁸ See, for example, Clarke (1705).

them to find that they are not altogether unworthy of regard, and that however their past conduct may be censured, their present disposition is at least approved of, and is perhaps sufficient to compensate the other, at least to maintain them in some degree of esteem with their friend. A numerous and artful clergy had, in those times of superstition, insinuated themselves into the confidence of almost every private family.... and hence the origin of books of casuistry. (TMS 7.4.17, 333–4)

To oversimplify somewhat: in normal circumstances one must find some compensating mechanism to rid oneself from psychological distress. As we have seen, on Smith's anti-clerical interpretation,²⁹ the clergy's development of casuistry exists to supply this need.³⁰

It is tempting to think that according to Smith his (and our) times are no longer superstitious – that this is simply run-of-the-mill Enlightenment rhetoric of progress. But if that were the case, it would make no sense for Smith to introduce the language of the “piacular” into his treatment of that species of shame involved in being an involuntary cause of harm. After all, from the point of view of reason, propitiatory sacrifice and a focus on purification just seem highly superstitious. Moreover, while I suspect that many reasonable readers would be willing to agree with Smith that some response is appropriate to being an unwilling cause of harm (for instance, a response along the lines of, say, Williams' agent-regret), few would want to follow Smith into the language of pollution and shame. Smith himself signals that he is at least aware that on this score he is hardly being a “modern”.³¹ He describes the feeling as “what the ancients called, piacular.” *These ancients are pagans of the “heathen religion.”*³² In fact, in Smith's *Lectures on Jurisprudence*, “superstition” is basically treated as synonymous with attempts at bribing deities (with prayer or sacrifice) for some favor to alleviate one's fear.³³

²⁹ See also Fleischacker (2005: 71).

³⁰ Smith's treatment of the way scientific inquiry responds to the painful feeling of wonder is another instance of this general psychological mechanism (see Schliesser 2005). In the “Astronomy” Smith alludes to the idea that science just is a very refined species of superstition, but that topic must be explored elsewhere. I thank Leandro Stieben for conversation on this point.

³¹ For further on this theme more generally, see Berns (1994), Montes (2004), Hanley (2006). Cf. Vivenza (2002).

³² The piacular is also compatible with Judeo-Christian ideas (see below). However, the word “piacular” and its variants are extremely rarely used in extant Latin works (and entirely absent in Roman, Christian sources), but a Perseus search reveals the term occurs quite often in Livy, which also seems to be Smith's source here (and which as we know from Smith's Rhetoric, he knew well), and also in Lucan's *Pharsalia*, a text he would have (re-)read in light of the epigraph Book 3 of the *Treatise*.

³³ Smith treats the slaves' inability to engage in sacrificial bribing of the Gods as one of the main causes of the growth of Christianity in *Lectures on Jurisprudence*: “I observed before that superstitious fears and terrors increase always with the precariousness and uncertainty of the manner of life people are engaged in, and that without any regard to their religion.... Slaves were of all others the most dependent and uncertain of their subsistence. Their lives, their liberty, and property were intirely at the mercy of the caprice and whim of another—It was therefore very hard that they who stood most in need of some consolation in this way should be intirely debarred from all religious societies, {which might at least sooth their superstitious dreads}. The gods then were alltogether

Now, if Smith had just used the language of shame and avoided the language of sacrifice, it would not be clear that the third aspect of atonement I called attention to above (recall (iii): acknowledging that one is, in some sense, marked or polluted) is all that significant. But Smith's treatment of the piacular emphasizes sacrifice. Thus, Smith's whole discussion of the piacular swerves uncomfortably towards the position that there are circumstances in which piacular shame is justified while being akin to superstition (again, I ignore the complex question of whether this is a moral feeling). Presumably, the only reason why the piacular is not ultimately classified as an instance of superstition is because the wise and virtuous are not motivated in their atonement by that fear which is held to be the essential core of superstition (see also Hume's "Of Superstition and Enthusiasm", *EMPL*, 73–7).

The significance of the nature of superstition in the piacular becomes clear when we reflect on a closely related case in Smith:

Gratitude and resentment ... are excited by inanimated, as well as by animated objects. We are angry, for a moment, even at the stone that hurts us. A child beats it, a dog barks at it, a choleric man is apt to curse it. The least reflection, indeed, corrects this sentiment, and we soon become sensible, that what has no feeling is a very improper object of revenge. When the mischief, however, is very great, the object which caused it becomes disagreeable to us ever after, and we take pleasure to burn or destroy it. We should treat, in this manner, the instrument which had accidentally been the cause of the death of a friend, and we should often think ourselves guilty of a sort of inhumanity, if we neglected to vent this absurd sort of vengeance upon it. (TMS 2.3.1.1, 94)

A stone that hurts us is also an involuntary cause of harm. In this sense, it is structurally identical to cases of the piacular. Reactive attitudes toward a stone that harms us are natural and even humane, and yet they are "improper" and "absurd." While Smith clearly could have thought that piacular shame and the feeling of pollution would also be absurd – as many modern moral philosophers are wont to think, opting instead for the more sensible notion of agent-regret – Smith suggests that the piacular feeling is perhaps not so absurd. According to him, the piacular feeling of shame teaches us "to reverence the happiness" of our "brethren." (TMS 2.3.3.4, 106).³⁴ Thus, paradoxically, it is when we are most like a stone in the inevitable causal chain of nature that we are most in a position to be taught the proper attitude toward our common humanity (which is really quite different from a stone).³⁵ A further important consequence of this, according to Smith, is that we are forced to adopt a kind of do-no-harm principle: "to tremble lest [we] should, even unknowingly, do any thing that can hurt" our "fellow brethren." (TMS 2.3.3.4, 106)

locall or tutelary; they did not conceive any god that was equally favourable to the prayers of all.... Besides, the deities then could never be addressed empty handed; who ever had any request to ask of them must introduce it with a present. This also intirely debarred the slaves from religious offices as they had nothing of their own to offer; all they possessed was their masters.... This it was which made all religions which taught the being of one supreme and universall god, who presided over all, be so greedily receivd by this order of men." (LJ, Tuesday. February 15th. 1763.)

³⁴ According to Smith, we can feel reference for parents and even more for our children, for God, the laws, duty, political leaders, etc. Smith returns to the significance of our fellow-feeling with our brethren throughout TMS.

³⁵ As an aside, here Smith is at his least Spinozistic because he seems to reject pan-psychism.

To avoid confusion, I am not claiming that Smith thinks that propitiatory sacrifice is always superstitious. For, as he wrote in the original ending of edition 1–5 (though this was removed in the sixth edition):

The doctrines of revelation coincide, in every respect, with those original anticipations of nature; and, as they teach us how little we can depend upon the imperfection of our own virtue, so they show us, at the same time, that the most powerful intercession has been made, and that the most dreadful atonement has been paid for our manifold transgressions and iniquities. (TMS 2.2.3, 92)

Here, Smith is self-consciously echoing the *New Testament*, e.g., “Look, the Lamb of God, who takes away the sin of the world!” (John 1:29) and “This cup is the new covenant in my blood, which is poured out for you” (Luke 22:20:).

Even if for the sake of argument we assume that Smith was entirely sincere in these lines (and ignore the Epicurean treatment of Fortune above), there are two things to note about the case of the ordinary piacular. First, Smith provides no hint that he is assimilating the piacular to an instance of the imitation of Christ; second, Christ’s sacrifice does not eradicate the need to purify the person that does unwilling harm.³⁶

Even if the foregoing is correct, it is still somewhat unclear why Smith harkened back to a heathen concept that, at the very least, still flirts with superstition in the way it is associated with ideas that emphasize propitiatory sacrifice, purification and pollution. Moreover, given that we rarely are the unwilling causes of the deaths of innocents and we ordinarily tend to think of ourselves as agents, there seems to be no way to make sense of the piacular. Perhaps it must be conceded that there is something fundamentally unreasonable about it.

However, Smith claims that the pay-off of these not quite reasonable ideas is that “the happiness of every innocent man is, in the same manner, rendered holy, consecrated, and hedged round against the approach of every other man.” Perhaps the insight here is a simple one, namely, that we require some of these not entirely reasonable notions in order to activate feelings that will make us care in the right way for “the happiness of every innocent man”. On this point, Smith’s writings reveal a concern with the fact that even humane philosophers, who ought to know better, sometimes do not consider the happiness of all. As he writes regarding Plato’s attitude toward infanticide: “The humane Plato...with all that love of mankind which seems to animate all his writings, no where marks this practice with disapprobation.” (TMS 5.2.15, 210) Given that Plato was untouched by doctrines of original sin, Plato ought to have seen infants as entirely innocent, but he fails to show sufficient regard for them. In particular, Smith recognizes that neither our ordinary sentiments nor even the “love of mankind” can ensure that we take the happiness of all innocent people into consideration. As Smith writes in a passage added to the second edition of TMS:

It is not the love of our neighbour, it is not the love of mankind, which upon many occasions prompts us to the practice of those divine virtues. It is a stronger love, a more powerful affection, which generally takes place upon such occasions; the love of what is honourable and noble, of the grandeur, and dignity, and superiority of our own characters. (TMS 3.3.5.4, 137)

³⁶ This second point may be so, as Ryan Hanley suggested, because Christ’s remission of sin need not alleviate guilty feelings.

A willingness to care about the happiness of all requires divine virtues, including those that are in some sense a kind of superior or highly elevated self-love.³⁷ The highest form of the piacular is exhibited only in those capable of such divine virtues. This leaves open, of course, why Smith thought such people, in particular, require that others are rendered holy to them.³⁸

10.7 Conclusion

As we have seen, according to Smith, in cases of inanimate instruments that harm us, sometimes nature and reason are not in harmony with each other. A well-known instance of this is his claim: “That kings are the servants of the people, to be obeyed, resisted, deposed, or punished, as the public conveniency may require, is the doctrine of reason and philosophy; but it is not the doctrine of Nature.” (TMS 1.3.3, 53) In this context, Smith seems to side with nature (see also 3.3.8, 139). Instances of the piacular then are significant because in them nature itself is revealed as – to use a Smithian word – irregular³⁹: the family and friends of those harmed feel a natural but nonetheless unjust resentment. And while Smith implies that the majority of us who find ourselves in the unlucky position of being unwilling causes will likely flee to casuists (clerical or secular) who can supply us with excuses to avoid carrying around our discomfort and offer us ways to atone, a few, more noble and sensitive souls – those with the soundest judgments⁴⁰ – will remain marked forever. Smith seems to have believed that reflection on cases of the piacular can help us appreciate the moral significance of all.

I read Smith as embracing the idea that aspects of superstition play an essential and sometimes positive role in the workings of our moral life. One might think that my evidence for this claim is slender given that it hangs on the interpretation of three paragraphs added to TMS. But consider also the following passage:

If the injured should perish in the quarrel, we not only sympathize with the real resentment of his friends and relations, but with the imaginary resentment which in fancy we lend to the dead, who is no longer capable of feeling that or any other human sentiment. But as we put ourselves in his situation, as we enter, as it were, into his body, and in our imaginations, in some measure, animate anew the deformed and mangled carcass of the slain, when we bring home in this manner his case to our own bosoms, we feel upon this, as upon many other occasions, an emotion which the person principally concerned is incapable of feeling, and which yet we feel by an illusive sympathy with him. The sympathetic tears which we shed for that immense and irretrievable loss, which in our fancy he appears to have sustained, seem to be but a small part of the duty which we owe him. The injury which he has

³⁷ See Hanley (2009).

³⁸ This is not the place to explore Smith’s complex treatment of magnanimity. See [Corsa \(ms\)](#).

³⁹ To avoid confusion, I do not mean to imply laws of nature are broken.

⁴⁰ Smith is sometimes treated as a defender of ordinary common sense, but it is worth noting that he also wrote: “I shall give an instance in things of a very frivolous nature, because in them the judgments of mankind are less apt to be perverted by wrong systems.” (TMS 1.1.3.3, 17)

suffered demands, we think, a principal part of our attention. We feel that resentment which we imagine he ought to feel, and which he would feel, if in his cold and lifeless body there remained any consciousness of what passes upon earth. His blood, we think, calls aloud for vengeance. The very ashes of the dead seem to be disturbed at the thought that his injuries are to pass unrevenged. The horrors which are supposed to haunt the bed of the murderer, the ghosts which, superstition imagines, rise from their graves to demand vengeance upon those who brought them to an untimely end, all take their origin from this natural sympathy with the imaginary resentment of the slain. (TMS 2.1.2.4, 71)

The enlivening of the idea part of the sympathetic process can naturally and justifiably draw on counterfactual imagery that by Smith's light is "superstition." This superstition is the product of "Nature" which has "stamped upon the human heart, in the strongest and indelible characters, an immediate and instinctive approbation of the sacred and necessary law of retaliation." (TMS 2.1.2.4, 71) Or, to put it simply, at the heart of those natural mechanisms and reactive attitudes which form the foundations of moral thinking and justice we encounter superstition.⁴¹ Of course, if we lack fear we will not ourselves be superstitious even if we engage in superstition.

Elsewhere in treating Smith's historical understanding of property rights, I claimed that for Smith accounts of the naturalistic development of social institutions of property and their *articulation* in our practices and theorizing are partially constitutive of the normative force of these institutions. Of course, this account is not the sole authority of the normative force of these institutions. Smith himself emphasizes that for most of us, this force is for the most part derived from long-standing moral rules, whose authority, in turn, is derived from other (divine) authorities. But in a world of Enlightenment, any appeal to authority – whether reason or revelation – is always open to further investigation. In particular, Smith recognizes that Enlightenment imperative which demands non-miraculous, causal explanations for our practices.⁴² Such explanations teach us something about the social conditions of possibility of our institutions and practices. The historical account enables a moral theory in which explanation and justification are mutually reinforcing for affective beings self-conscious of their status as intellectual animals. Yet, we also see that sometimes this investigation reveals the "absurd," "improper" and even "superstitious" nature of mankind. While an uncharitable reader might think this is part and parcel of Enlightenment elitism, as we have seen, Smith's approach includes the wise and virtuous among those that draw on superstitious elements.⁴³ There is, thus, no reason for thinking that the "precise and distinct measure" that can be found "in the sympathetic feelings of the impartial and well-informed spectator" would be exempt from this.

⁴¹ Of course, not just any kind of superstition – presumably only those views that help fortify our sense of justice. This has resemblance to Smith's views on how morality constrains and regulates theology. See Schliesser (2008). I thank Chad Flanders for discussion on this point.

⁴² Contrast Kant's embrace of some historical taboos; see Kant (1797: 318–9, 339–40).

⁴³ Because they lack fear I hesitate to call them "superstitious," however they do rely on numerous superstitious elements. I thank Andrew Corsa for discussion on this point.

Acknowledgements I thank James Harris for his encouragement. I am grateful to Martin Lenz and Anik Waldow for their invitation and generous feedback. Audiences in Berlin and at Notre Dame, where Ryan Nichols, Sam Newlands, Aaron Garrett, and especially Terence Cuneo offered me lots of critical and helpful feedback, were extremely generous with suggestions. I am grateful to Chad Flanders, Ryan Hanley, and Andrew Corsa for important criticism on the final draft.

References

- Berns, Laurence. 1994. Aristotle and Adam Smith on justice: Cooperation between ancients and moderns? *The Review of Metaphysics* 48(1): 71–90.
- Berry, Christopher. 1974. Adam Smith's considerations on language. *Journal of the History of Ideas* 35(1): 130–138.
- Berry, Christopher. 2006. Smith and science. In *Cambridge companion to Adam Smith*, ed. Knud Haakonssen. Cambridge: Cambridge University Press.
- Brissenden, R.F. 1969. Authority, guilt, and anxiety in the theory of moral sentiments. *Texas Studies in Literature and Language* 11(2): 945–962.
- Clark, Andy. 2012. (in press). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*.
- Clarke, Samuel. 1705. *A demonstration of the being and attributes of god*. London: J. Knapton.
- Corsa, Andrew (ms) David Hume: Adam Smith's great-souled man.
- Debes, Remy. 2007. Humanity, sympathy and the puzzle of Hume's second enquiry. *British Journal for the History of Philosophy* 15(1): 27–57.
- Flanders, Chad. 2006. "This irregularity of sentiment": Adam Smith on moral luck. In *New voices on Adam Smith*, ed. Leonidas Montes and Eric Schliesser. London: Routledge.
- Fleischacker, Samuel. 1999. *A third concept of liberty: Judgment and freedom in Kant and Adam Smith*. Princeton: Princeton University Press.
- Fleischacker, Samuel. 2005. *On Adam Smith's "Wealth of Nations": A philosophical companion*. Princeton: Princeton University Press.
- Garrett, Aaron. 2004. Adam Smith on moral luck. In *Adam Smith als moralphilosoph*, ed. Christel Fricke. Berlin: de Gruyter.
- Gibbard, Allan. 1990. *Wise choices, apt feelings: A theory of normative judgment*. Cambridge, MA: Harvard University Press.
- Hankins, Keith. (ms) Adam Smith on moral luck and the wide scope of moral responsibility.
- Hanley, Ryan P. 2006. Adam Smith, Aristotle and virtue ethics. In *New voices on Adam Smith*, ed. L. Montes and E. Schliesser. London: Routledge. Chapter 2.
- Hanley, Ryan P. 2009. *Adam Smith and the character of virtue*. Cambridge: Cambridge University Press.
- Hanley, Ryan P. 2010. Scepticism and naturalism in Adam Smith. In *The Adam Smith review*, ed. V. Brown and S. Fleischacker, 198–212.
- Hanley, Ryan P. 2011. David Hume and the "Politics of Hu-manity". *Political Theory* 39(2): 205–233.
- Hume, David. 1987. In *Essays moral, political, literary*, ed. Eugene F. Miller. Indianapolis: Liberty Fund.
- Hume, David. 2004. In *A treatise of human nature, volume 1: Texts*, ed. David Fate Norton and Mary J. Norton. Oxford: Oxford University Press.
- Kant, Immanuel. 1797 [1902]. *Die Metaphysik der Sitten*, der Akademie-Ausgabe Gesammelten Werke, vol. VI, Electronic edition, Berlin: Walter de Gruyter.
- Levy, David M. 1997. Adam Smith's rational choice linguistics. *Economic Inquiry* 35(3): 672–678.
- Montes, Leonidas. 2004. *Adam Smith in context a critical reassessment of some central components of his thought*. London: Palgrave Macmillan.

- Otteson, James. 2002. Adam Smith's first market: The development of language. *History of Philosophy Quarterly* 19(1): 65–68.
- Pack, Spencer, and Eric Schliesser. 2006. Adam Smith's 'Humean' criticism of Hume's account of origin of justice. *Journal of the History of Philosophy* 44(1): 47–63.
- Russell, Paul. 1999. Smith on moral sentiment and moral luck. *History of Philosophy Quarterly* 16: 37–58.
- Schliesser, Eric. 2005. Realism in the face of scientific revolutions: Adam Smith on Newton's "Proof" of copernicanism. *British Journal for the History of Philosophy* 13(4): 697–732.
- Schliesser, Eric. 2006. Articulating practices as reasons. *Adam Smith Review* 2: 69–97.
- Schliesser, Eric. 2008. Review of D. D. Raphael the impartial spectator: Adam Smith's moral philosophy and Leonidas Montes Adam Smith in context. *Ethics* 118: 569–575.
- Schliesser, Eric. 2011. Reading Adam Smith after Darwin: On the evolution of propensities, institutions, and sentiments. *Journal of Economics and Behaviour* 77(1): 14–22.
- Schliesser, Eric. Forthcoming. Counter-factual causal reasoning in Smithian sympathy. *Revue internationale de philosophie*.
- Smith, Adam. 1982a. *The theory of moral sentiments*. Vol. I of the Glasgow edition of the works and correspondence of Adam Smith, ed. Raphael D. D., and A. L. Macfie. Indianapolis: Liberty Fund.
- Smith, Adam. 1982b. *Essays on philosophical subjects*. Vol. III of the Glasgow edition of the works and correspondence of Adam Smith, ed. Wightman W. P. D., and J. C. Bryce. Indianapolis: Liberty Fund.
- Smith, Adam. 1985. *Lectures on rhetoric and belles lettres*. Vol. IV of the Glasgow edition of the works and correspondence of Adam Smith, ed. Bryce, J. C. Indianapolis: Liberty Fund.
- Taylor, Jacqueline. 2006. Humean humanity versus hate. In *The practice of virtue*, ed. Jennifer Welchman, 182–203. Indianapolis: Hackett.
- Vivenza, Gloria. 2002. *Adam Smith and the classics: The classical heritage in Adam Smith's thought*. Oxford: Oxford University Press.
- Williams, Bernard. 1976. Moral luck. *Proceedings of the Aristotelian Society, Supplementary Volumes* 50: 115–135. &137–151.

Chapter 11

Explaining and Describing: Panpsychism and Deep Ecology

Michael Hampe

11.1 Outlining the Topic

In the following, the deep ecology project, particularly as established by the Norwegian philosopher Arne Naess, will be subjected to philosophical examination. Deep ecology is of philosophical interest for *three* reasons: firstly, in terms of *method*, as it is usually classed and criticised as a form of environmental ethics; secondly in terms of content, due to its criticism of *anthropocentrism*; and thirdly, again methodologically, because it aims to bring about a fundamental *change in our civilisation* through philosophical theories (of perception and of reality), or, more specifically, through texts which constitute the theoretical part of what Naess refers to as a ‘movement’.¹

As to the first point, the classification of deep ecology into environmental ethics: for pedagogical and organisational reasons, there exists a ‘division of labour’ in academic philosophy. Philosophy is divided into different areas in the same way that physics is subdivided into theoretical and experimental physics, biology into molecular biology and genetics, and chemistry into organic and inorganic chemistry. There is theoretical and practical philosophy, and practical philosophy in turn is subdivided into political philosophy, meta-ethics, normative ethics, applied ethics and sometimes even (in Zurich) into methods of applied ethics. This is useful for structuring teaching in the individual philosophical disciplines and for developing specialised knowledge within the faculty. However, if one considers philosophical ‘schools’ and *works* such as Charles Sanders Peirce and William James’ pragmatism, Max Horkheimer and Theodor W. Adorno’s critical theory, or Spinoza’s *Ethica*,

¹ See, for instance, the title of Naess’s programmatic work, “The shallow and the deep, long-range ecology movement. A summary.” Naess (1973: 95–100).

M. Hampe (✉)
ETH Zürich, Zurich, Switzerland
e-mail: hampe@phil.gess.ethz.ch

such divisions tend to cause a hindrance rather than aiding the understanding of these works. Unlike Millian utilitarianism, Charles Sanders Peirce's and William James' pragmatism is not just a line of thought in practical philosophy, although it does call for adherence to the principle that no conceptual distinctions made in philosophy should be accepted that do not correspond to some distinction in living practice.² As Horkheimer's critical theory itself calls into question the categories of "the better", "the useful", etc., points of orientation for individual and collective action, and thereby also the foundations of most ethical investigations in philosophy, one would hardly want to simply classify the theory as practical or political philosophy either.³ And although Spinoza's opus magnum bears the title 'ethica', it cannot simply be regarded as pertaining to practical philosophy. After all, historically it has been received – correctly – first and foremost as an *ontology* and *philosophy of mind*. Briefly put, more creative and original philosophical ideas and works will not adhere to academic divisions; rather they will dispel or cross them. As this paper will argue, deep ecology is misunderstood because it is not analysed primarily as a *critique of civilisation* (which is not a philosophical 'discipline') but as *applied ethics*. This is probably due to the inability on the part of the respective reception to suspend the established methodological classification in philosophy for the case in question, which constitutes a web of thoughts that will not fit the pattern of so-called 'classical texts'.

On the second point, the criticism of anthropocentrism: while it seems to me that, intuitively, deep ecology's criticism of the anthropocentrism in modern western thought and behaviour is certainly correct, it does seem absurd at first glance. From which perspective, it will be asked, how are people to *view* the world and *assess* its events and their own actions, if not from that of a human being?⁴ Have we not painstakingly learnt ever since the seventeenth century that we do not have access to a *divine* perspective and can, as it were, not dissociate ourselves from our being human? Is inviting people to learn to perceive the world from the perspective of other, non-human species, as deep ecology does, not a relapse into pre-enlightenment thinking? If, following Bertrand Russell, the ability to adopt a perspective on something is made the criterion for *subjectivity* and *the mental*,⁵ then, by suggesting that people ought to adopt the perspective of non-human natural life, deep ecology is implying that not only humans are capable of subjectivity and of being in mental states. But how can we ever be certain of this? Given the mere possibility of solipsism or the problem of other minds, it is difficult to produce an argument for the claim that we really *know* that not only we as individuals, but *other people*, too, have a perspective on the world and are in mental states, and are not simply machines or

² Cf. James (1982, 442 f).

³ Cf. Horkheimer (2002, 181).

⁴ The anthropocentrisms taken together here have also been distinguished into epistemic and moral anthropocentrism. Cf. Krebs (1996: 351).

⁵ Cf. Russell (1921, 105).

zombies.⁶ On the other hand, beyond the method of doubt, to claim that one doesn't know whether one's father, daughter, dog or the horse one is riding has a perspective on the world seems to be an *absurd scepticism*. Is it not a completely reasonable request to ask a rider to put himself into his horse's position? But to which animals do we naturally ascribe subjectivity and mental states, and to which do we not? We naturally consider a chimpanzee, a horse or a dog to be subjective; but what about a lobster when it tries to escape from the pot? Who would be capable of boiling a live rabbit in the same fashion?⁷ Our relationship with the world makes subjectivity extend beyond the frayed edges of humanity, so that pets and some wildlife, such as elephants and dolphins, are regarded as being subjective. But beyond this, we lose the ability to consider other natural creatures as subjective counterparts. Deep ecology wants to change our relationship with the world in such a way as to make us view *more* non-human forms of life as subjective counterparts than in the past.

We expect children to develop the ability to literally take the spatial perspective of another person at a certain age, something they are still unable to do at the age of 3 or 4 years.⁸ It is hardly a sign of moral maturity (although perhaps philosophically sophisticated) to doubt other people's subjectivity by taking a solipsistic stance. Following Stanley Cavell, it can be argued that by taking this sceptical attitude towards the subjectivity of others one is cut off from a certain relationship with the world in which others are naturally experienced as subjects.⁹ What deep ecologists evidently demand when they say that humans must learn to take the perspective of other, non-human forms of life is a collective process of maturation in this very sense. Just as racism was the collective inability to adopt the perspective of other people and had to be overcome by a process of sensitisation.

In the eyes of deep ecologists, anthropocentric speciesism, likewise, is a collective lack of sensitivity and maturity. The positive assessment of collective processes as "maturing" is reminiscent of concepts from the *progress of civilisation*. Deep ecology, it seems, aims at achieving something akin to promoting the progress of civilisation by philosophical means when it recommends regarding more forms of life as being subjective counterparts.

This takes us to the third point, the idea of a *change of civilisation*. Deep ecology does not believe that the process of maturing referred to in the previous paragraph can be realised by working out a utilitarian or deontological applied ethics. It is concerned with something more fundamental which supposedly underlies these ethical principles: the way in which human beings *experience* non-human nature. Now, there can be no doubt that the way in which people *experience* the world and non-human nature, is influenced by the patterns in the ways people *live*, which can

⁶This possibility was first considered in Descartes' Second Meditation. Cf. Descartes (1973, 32).

⁷Wallace (2006, 245).

⁸On the subject of the psychology of a change of perspective, refer to the classical experiments by Jean Piaget and Bärbel Inhelder in Piaget and Inhelder (1956) and, more recently, Huttenlocher and Newcombe (2003).

⁹Cf. Cavell (1979, part Four): "Skepticism and the Problem of Others".

be summed up by the concepts of “civilisation”, “world view”, “form of life” etc. (each of which seems to signify something different yet very fundamental and collective). And without doubt, philosophy has influenced the development of these patterns at times – one only needs to think of the impact of Karl Marx’s theory. But it is debatable whether philosophy is actually able to *control* processes of civilisation by developing theories.¹⁰ Even if, in some projects, it may seem as if the history of western civilisation were an epiphenomenon of the history of philosophy (Heidegger’s history of the forgetfulness of being can be interpreted in this way), it is equally possible for philosophy itself to be viewed as an *expression* of cultural processes or of processes of civilisation.¹¹

Thus, with its criticism of anthropocentrism, Naess’ deep ecology constitutes a *critique of critical enlightenment thinking* because it calls into question the privilege of the third-person perspective in establishing knowledge of nature and wants to regard elements of living nature as quasi-personal counterparts. Scepticism of the communicated structures of the inner worlds of other people, and, indeed, their very existence, is an integral part of modern enlightenment thinking. Deep ecology’s criticism of anthropocentrism is therefore a fundamental criticism of the Enlightenment; hence it is sometimes even interpreted as being anti-Enlightenment and as constituting a return to dangerously naïve ideas. It is true that criticism of the Enlightenment is not new as such. Indeed, it appeared in philosophy parallel to Enlightenment thinking and became central to philosophy in the twentieth century in authors like Adorno, Horkheimer and Heidegger. However, the appeal to regard non-human life as a ‘You’ can also be considered a “relapse” into mythical thinking.¹² But this is certainly not Naess’ intention. He believes that a new theory of perception and an ontology which operate with the concept of gestalt can contribute to bringing about a new human state of civilisation. This claim does, however, seem problematic. Therefore, it is necessary to take a closer look at the collective patterns referred to as “civilisation”, “world view”, “form of life”, etc., which were mentioned briefly earlier on, and to shed more light on their relationship to philosophy in order to better understand what practical relevance the development of philosophical theories such as deep ecology might have in the first place.

11.2 Applied Ethics and Deep Ecology

In philosophy, the actions of human beings in landscapes and towards plants and animals can be considered in two ways. Insofar as actions are assessed, the first approach will translate into a form of *applied ethics*. Thus, as one may ask what *rights and obligations* people have in their dealings with *one another*. Another consideration could be whether landscapes, plants and animals have rights and whether people have

¹⁰ I would like to thank Maria-Sibylla Lotter for the thought underlying this question.

¹¹ This is what Sigmund Freud seems to imply if his remarks on ethics and religion are also brought to bear on philosophy. Cf. Freud, (1930, 503).

¹² Cf. Theobald (2003, 69–86). Theobald criticises the conflation of mythical and rational (in the sense of enlightenment critique) elements in deep ecology.

obligations towards these non-human forms of life. Furthermore, one might also ask whether, and if so, to which extent it is *useful* or not to protect non-human nature. This means that, following Kant, the actions of human beings in landscapes as well as towards plants and animals can be studied in terms of a *deontological* or a *utilitarian* assessment; hence principles that are also used in other areas of ethics can equally be *applied* to this sphere of reality, which is affected by human actions. Loosely speaking, what is known as environmental ethics is divided into these two forms of applied ethics when asking, for example, whether animals such as dolphins or chimpanzees have rights and how it would be possible to represent these in the human community, or when studying whether it is more useful to preserve a wetland or drain the area and approve development in a region where there is a housing shortage.¹³

The other possibility is to interpret the way in which human beings behave in landscapes and towards plants and animals as indicative of their *experience* or *understanding* of this non-human nature. These ways of experiencing and understanding natural forms of life may be accessible both in writing or orally, explicated or provided implicitly by people, for example in certain forms of ritual behaviour. They are contingent upon upbringing, language, theory and history. These ways of experiencing and understanding can be referred to as the *ontologies* of non-human forms of nature that exist in certain cultures. In this sense, the implicit ontology of plants and animals of the Jivaro in the Amazon in the fifteenth century differs from that of today's North Americans and Europeans.

Responding to the ontological question of what non-human life forms *actually are* for certain people also means adopting a view on what *human beings* actually are, since the very question draws a distinction between people and non-human nature. Ontological beliefs therefore go hand in hand with anthropological beliefs. Moreover, beliefs about the nature of natural beings influence the *assessments* of actions towards these forms of natural life. He who believes that insects are unable to feel pain due to the structure of their nervous system will assess the slow killing of insects by people differently to someone who believes that they very much do feel pain. Ontological and normative beliefs are therefore inseparable.

Essential parts of what is known as “deep ecology” also deal with the question of what non-human forms of life are, the intention being, however, to assess and change human behaviour towards these beings. The fact that this is deep ecology's declared objective (a fact frequently overlooked in debates on environmental ethics) is illustrated by the following statement from one of its founders, Arne Naess:

I am not very much interested in ethics or morals. I'm interested in how we experience the world... If deep ecology is deep it must relate to our fundamental beliefs, not just to ethics. Ethics follow from how we experience the world. If you articulate your experience then it can be a philosophy or religion (Naess 1989, 20).¹⁴

¹³For a more precise analysis of the variants of environmental ethical theory, see Birnbacher (2001, 103–139), and Krebs (1996, 350–352).

¹⁴See also the following statements from Naess: “But one's ethics in environmental questions are based largely on how one sees reality” (Naess 1989, 66). And: “It is, I think, important in the philosophy of environmentalism to move from ethics to ontology and back. Clarification of differences in ontology may contribute significantly to the clarification of different policies and their ethical basis.” (op. cit., 67)

Environmental ethicists also endeavour to change human behaviour or environmental policy. If environmental ethicists are concerned with deontology, they can attempt to clarify for people precisely which rights and obligations they must keep in mind in their actions towards animals and plants or, if they approach the issue from a utilitarian perspective, they can try to clarify which utilitarian criteria must be observed, repeatedly explicating these rights, obligations and utilitarian criteria in philosophical publications and political debates.¹⁵

By contrast, the effort to alter the behaviour of humans towards the environment presents itself as a far more difficult task for deep ecologists.¹⁶ This is because the deep ecologist must try to change *current views* of what certain natural beings *are*, the *very experiences* with nature, that serve as the basis for the ecologists' judgement that certain actions are wrong, and thus try to influence the implicit ontology. Or, to put it more simply, deep ecologists must endeavour to influence the way in which people experience the world in order to be able to alter human behaviour. They wish to alter the ontology of nature by changing the modes of experience and interpretation and alter the norms for actions by changing this ontology. The objective here is to create a *biosphere egalitarianism* in which no form of life is singled out as being more valuable than another, and in which people develop an understanding "from within" of other forms of life, an attitude which is usually (or *should be*) adopted only towards other human *persons*.¹⁷ The same right to live and develop should be established for all living creatures as an intuitively insightful axiom of thought and action. However, these axioms can indeed only be understood and acknowledged by human beings. In this respect, the human-animal relationship remains asymmetrical in deep ecology as well.¹⁸

Since in deep ecology axioms and intuitions are not justified, but should be self-evident, they form the *foundation* of a process of justification. Deep ecology is therefore concerned with changing the foundations of human thought about, and actions within, natural contexts. Although these foundations are not *justifiable*, they are, however, *motivated* by experiences. Experiences cannot serve as justifications of anything, because they are not propositions. Yet, although the axiom "avoid pain" is not justifiable, it is well *motivated* for those who have experienced pain. Someone who does not know pain will hardly start from this axiom. This is why different *horizons of experience* are decisive for different intuitive foundations of justifications and actions. For this reason deep ecology is neither a form of applied ethics, nor is it a form of ethics based on metaphysics, as it is sometimes claimed and objected.¹⁹

¹⁵ Cf. Regan (1997), Wolf (1997), and Frey (1997). Also see Ott (1997, ch. 7) and Leist (1996, 388–456). For utilitarianism, see for example Peter Singer's influential Singer (1979, ch. III) and Wolff (1992).

¹⁶ For the objectives of deep ecology's changes to civilisation, see the programmatic text by Arne Naess entitled "The shallow and the deep, long-range ecology movement. A summary" (Naess 1973, 95–100).

¹⁷ See Naess (1973, 95 f).

¹⁸ I would like to thank Jonas Lüscher for this important information.

¹⁹ For example, Pfordten (1996, 114–120).

Rather, it is a *sui generis* project which is therefore often also referred to as ecophilosophy or ecosophy or criticised for being mythical and religious and thus irrational.²⁰

Clearly, creating awareness of established ethical deontological or utilitarian principles and applying these principles to a specific area is an easier task (if also perhaps very laborious in its details) than the attempt to alter people's horizons of experience, that is, their *world views*.²¹ Certainly, one may wonder *if philosophy is at all able to alter horizons of experience or world views*. If we turn to the writings of Hobbes, Descartes and Spinoza, or Marx and Heidegger, it is obvious that philosophy has deemed itself capable of this time and time again. However, the interesting question here is whether philosophy has overestimated its power.

11.3 Ontology and Contradiction

It is an exaggeration to speak of “ontologies” – as we just did – when trying to collect the basic beliefs of humans by asking them what natural beings are, and what they themselves actually are, and whether they believe that they have an immortal soul, plants have feelings and animals have life plans. After all, the way in which these basic beliefs are conceptualised are seldom as systematically structured as Book Zeta in Aristotle's *Metaphysics* and his *Peri Psyche*, or the first part of Spinoza's *Ethica*, Christian Wolff's *Prima Philosophia* or the second and third chapters of Whitehead's *Process and Reality*; all texts which can be read as “ontologies”.

The behaviour and ways of speaking common in certain cultural areas are not, however, organised by philosophers; rather, these expressions of cultural life are self-organised and uncontrolled, while the law of non-contradiction is evidently not constitutive of this self-organisation.²² Human beings cope with life as individuals and as collectives, despite the fact that their thoughts and, more often, their assessments and actions are *contradictory*. Yet, like Spinoza, Naess also believes that contradictions between fundamental intuitions and behavioural maxims cause suffering on the part of the actor. Nevertheless, it is possible to live with such contradictions (albeit not well), even when suffering is part of this life. But still, webs of contradictory beliefs and possibly vague concepts should not be referred to as “ontologies”.

²⁰ For the distinction between different areas of applied ethics, see Naess (1997, 192), and Naess (1989, 87–103). For the criticism of deep ecology as mythical see Theobald, *op. cit.* (fn. 9).

²¹ Don E. Marietta speaks of “world view” in a similar context; see Marietta (1979, 200).

²² In connection with political philosophy, Raymond Geuss has pointed that people exist in contradictions, but that political philosophy mostly assumes that a “Cartesian subject” organises his beliefs and actions consistently, which is why its philosophy goes amiss of political reality (cf. Geuss 2008, 3 f.).

It is these supposed contradictions between experience, thought and action, along with the symptoms of ecological crisis that have existed in well-developed societies ever since the second half of the twentieth century, which provide Naess with the starting point for his critique of the current state of civilisation in North American and European societies. On the one hand, in everyday life human beings do indeed largely experience higher animals as individuals endowed with subjectivity and as having value in themselves. However, on the other hand, humans can neither conceive of animals in this way in the context of their objectivising sciences, nor are they able to let their actions conform to their perception of animals as creatures of value when faced with a food industry which serves and exploits the high meat consumption of millions of people.²³

Naess does not need to start from scratch when trying to render our concepts more consistent by changing the ways in which non-human beings are experienced, because we can already rely on experiences in which higher animals figure as subjective beings. However, given that these experiences are considered personal, they are *of merely anecdotal relevance* in a scientific culture which relies on *objective experiences* alone in the generation of knowledge; that is, experiences that can be reproduced in experiments and in which a scientific subject relates to a non-subjective object. And as long as science remains the authoritative source of information on nature in modern societies, so-called anecdotal experiences will continue to fail to have any political consequences. One of the objectives of deep ecology must therefore be to bring about a *societal upgrading* of everyday experiences of nature as opposed to the experiences we have under laboratory conditions. In a context in which laboratory experience favours the third-person perspective (“an object was observed”), this change entails the rehabilitation of a second-person perspective in our references to non-human nature (“you, dog, clearly want something to drink”).²⁴

The way in which human beings *live* and their *customary actions* are the principal factors that determine which trains of thought are replicated and deemed plausible even without the need for arguments and evidence to be produced. The same principle can also determine philosophical ontologies. On this conception, Aristotelian ontology emerges as the systematic conceptualization of a world experience and mode of life which have not yet been pervaded by an experimental sense. However, this way of conceiving of ontologies is not necessary. Philosophical ontologies can also be “revisionary”, i.e. represent perspectives that either *contradict* what is generally believed about that which exists or must be regarded as the conceptual foundation of everyday actions.²⁵ In this sense, although the ontology

²³ One could view Naess in the tradition of Rousseau here, for whom, according to Konersmann, a diagnosis of “a monstrous imbalance, the breach between form of life and form of thought” was central to the criticism of his contemporaries see Konersmann (2006a, b, 211, 235).

²⁴ With regard to the meaning of the second-person perspective in ethics and epistemology, which enters into the interpretation of deep ecology here, I owe much to my discussion with Lutz Wingert in Zurich.

²⁵ This distinction between revisionary ontologies and those supported by everyday actions follows Peter Strawson’s distinction between descriptive and revisionary metaphysics, although it does not coincide exactly with it (cf. Strawson 1959, 9–12).

of deep ecology can tie in with human beings' anecdotal everyday experiences of higher animals and the aesthetic experience of landscapes, it is revisionary because it wants to regard non-human natural beings as subjective. It is revisionary because it wants to extend the scope of subjectivity for non-human natural beings far beyond higher animals, and because it wants to stop the relegation of anecdotal experiences which cannot be operationalized scientifically in experiments by regarding these as culturally crucial experiences. This ultimately amounts to an attempt to shift the focus away from a culture dominated by experimental sciences in terms of its relationship with nature.

Now, thought and action are not simply interconnected in the human demand to live as one thinks (a demand which is seldom realised). Rather, the beliefs, language and habits in human beings' concrete lives outside of academia constantly reinforce and influence one another. Someone who regularly speaks of witches and conducts witch trials will be more easily convinced that witches exist than someone who does none of the above. And, vice versa, someone who is convinced that witches exist will be more likely to conduct a witch trial. A person will be familiar with the number pi if regularly calculating circumferences and the surface contents of spheres, whereas someone unaware of these kinds of practices will be unfamiliar with the number. Vice versa, someone who knows the number pi will be able to follow the corresponding calculation method. Similarly, people who are encouraged to separate their waste and read the associated pamphlets will know the word "sustainability" and become attuned to its meaning, whereas members of rainforest tribes will not readily understand this word. People who have eaten meat since their childhood will consider this practice as a normal activity and perhaps deem the associated production conditions "necessary" for the kind of life they consider normal. People who tend to eat meat only once a year, namely, on special occasions, will feel differently about this matter. The kind of thing that counts as "normal" and "feasible", for instance, when trying to establish an axiomatic basis for an argument, not only depends on a range of propositions, but also on the life that provides the context for uttering these propositions.

Revisionary philosophical world views, such as those conveyed in Spinoza's *Ethica*, Whitehead's *Process and Reality* or in chapters 3 and 7 of Naess' *Ecology, Community and Lifestyle*, are different kinds of "wholes" from those patterns of action and speaking habits characteristic of the "life" outside of academia and theory. Rather than being "philosophies" or theories, the latter contexts express something vaguely akin to a "world view" and "form of life". If, as is the case with deep ecology, one wishes to alter human beings' culture or civilisation in such a fundamental way that scientifically objectifying experiences of nature are no longer regarded as the paradigm of experience, one must ask whether philosophy can actually effectuate such change by impacting on us at the interface between actions, thought and speech. Further investigations of philosophical reflections on world views and related constructs are required in order to answer this question.

11.4 Worldviews in the Works of Jaspers, Wittgenstein's "Form of Life" and the Critique of Civilisation in Deep Ecology

There are no obvious systematic organisational principles or clearly identifiable authors structuring the way in which the potentially contradictory entities of thought, speech and action in life are organised as there are authors of philosophical systems. Karl Jaspers identified a worldview as a kind of knowledge that lies "beyond expertise"; something which not only consists of knowledge but also includes "evaluations, lifestyle, fate ... [and an] experienced ... hierarchy of values". For Jaspers, a worldview exists both "subjectively, as experience and force and mental attitude" and "objectively, as a physically shaped world."²⁶ In order to be able to separate the subjective from the objective aspect more effectively, Jaspers draws a distinction between "attitudes" ("Einstellungen"), that is, the subjective component of a worldview, and "pictures of the world" ("Weltbilder") as the objective parts, adding the following qualification: "Attitudes and world pictures are abstractions which isolate that which in fact co-exists ..."²⁷

For Jaspers, *evaluations* to which he refers as the "forces of life" ("Kräfte des Lebens") are of particular importance within this context. They exist as "something final": "There is no objective reason whatsoever," writes Jaspers, "as to why someone should make a value judgement. A human being does so as long as he lives; he can clarify his evaluations, formulate them, objectify them, but before he can do so they must be there and be experienced."²⁸ According to Jaspers, these "lived evaluations" reveal a hierarchy of values, which manifests itself in each of our "choices" in life. According to Jaspers, it is when this hierarchy is generalised by suggesting that others, too, ought to make specific choices that doctrines of life ("Lebenslehren") arise.²⁹ In this framework, the deep ecology movement could be interpreted as the attempt to establish a doctrine of life. Jaspers does, however, seem to doubt that it is possible to realise the aspirations of this doctrine of life: "However, the desire to establish absolutely a hierarchy of values based on some general evidence is hopeless, because any fixed hierarchy would prevent the possibility of new lived experiences of value relations."³⁰ In other words, the fundamental experiences of human beings cannot be homogenised; it is not possible to stabilise their horizons of experience over long periods of time.

The merging of the subjective and the objective – or the impossibility to differentiate between them in life – and the unjustifiability of the kind of evidence rooted in our everyday life not only inform Jaspers' concept of the worldview but also what

²⁶ Jaspers (1919/1960, 1).

²⁷ Jaspers (1919/1960, 219).

²⁸ Jaspers (1919/1960, 220).

²⁹ See Jaspers (1919/1960, 222).

³⁰ Jaspers (1919/1960, 226).

Wittgenstein calls a “form of life”. Wittgenstein himself uses this term very rarely and apparently with great unease, for instance, when in *On Certainty* he writes:

357. One might say ‘I know’ expresses *comfortable* certainty, not the certainty that is still struggling. 358. Now I would like to regard this certainty, not as something akin to hastiness or superficiality, but as a *form of life*. (That is very badly expressed and probably very badly thought as well.) 359. But that means I want to conceive it as something that lies beyond being justified or unjustified; as it were, as something animal.³¹

Wittgenstein does not speak of *evaluations* as Jaspers does; instead, he talks about *knowledge* and *certainty*. Nevertheless, he does in fact address the subject of “life” just like Jaspers does when he refers to the “animalist” not only in an *existential* but even *biological* sense. The fact that I know that I have never been to the moon is a certainty that correlates to with my “form of life”. It is something I do not need to justify by arguments and empirical studies. Whilst it could be the case that I have been to Wuppertal once and need to think about whether or not this is so, due to my specific “form of life” the question of whether I have been to the moon cannot possibly provoke the same train of thought as the one involved in my thinking about whether or not I have been to Wuppertal. In the life that I have experienced so far only a select few, American astronauts to be precise, have been to the moon, and, of course, I know that I am not an American astronaut. On the contrary, in a science fiction novel a different form of life could be depicted, one that no longer predominantly manifests itself on earth, but is characterised by travels to the moon or Mars, thus adding to the travels to Wuppertal or Augsburg that we undertake nowadays. Given such a form of life, it would be possible to consider whether or not one had already been to the moon. Nevertheless, in order to consider something I now think of as being “out of question” due to my current form of life, a great deal of very “fundamental” conditions would need to be different from the ones prevailing today.

Naess’ deep ecology also discusses the concept of “form of life” (or “way of life” and “lifestyle”), not in exactly the same sense as Wittgenstein, but not in a completely different sense either (Naess 1989: ch. 4). According to him, the idea that technical innovations and economic growth *must* exist is a matter of fact for us and something we assess positively in our current form of life. Yet, despite this, other forms of life are conceivable that do not consider this attitude as normal. Indeed, it is even said that it would be *desirable* for us to abandon our current form of life marked by our readiness to take these things for granted and to change to another one.³² The very title of the book in which this desire to change our current form of life is expressed contains the lexically relevant term: “Ecology, community and *lifestyle*.” Similar things can be said to motivate an understanding of the assumption of subjectivity in natural beings. Due to our lifestyle, we take it for granted that dogs, cats and horses are subjective beings with an inner world, while with respect to chicken, insects and fish the case is not so straightforward. If we interpreted these

³¹ Wittgenstein (1969, 31), my emphasis.

³² Naess (1989, 94, 110–122).

living beings as subjects, we would need to change many of the things we take for granted and would presumably also be forced to behave differently towards these living beings. Naess also refers to “normative systems”³³ in this context and the “established economic, social and technical framework”, a framework which the only operating ecological movement (i.e. the classic environmental ethics movement) attempts to change *bit by bit*. By contrast, the deep ecology movement is more radical insofar as it seeks to influence *civilisation as a whole*: “The aim of supporters of the deep ecology is not a slight reform of our present society, but a substantial reorientation of our whole civilization.”³⁴

11.5 Self-realisation, Spinoza and Panpsychism

Naess tries to realise this change of civilisation or our form of life through his philosophy (that parallels his activism), that is, by offering an account of ethical intuitions that stresses the need of a new panpsychic gestalt ontology, according to which perceptions emerge during a process of *gestalt development*. According to him, the distinction between a subject and an object, on the one hand, and that between facts and values, on the other, result from a process of abstraction that follows on from the actual processes of perception, and as such count as the products of post-hoc reflections about processes responsible for gestalt formation (while this reflection is itself another process of gestalt formation; see Naess (1989, 57–62). According to Naess, animistic mythological thinking that addresses mountains and rivers like people, deals directly with the subject of these kinds of “Gestalt” and does not assume that mountains or rivers have any other features apart from those we assign to them in our form of life.³⁵ Little of what is transferred from the mythological form of life or civilisation to our own involves the same objects being experienced in a different way, and just as little of what is transferred from our current form of life to the future form of life – the future form of life that deep ecology strives to achieve – will involve experiences of *the same* objects. Instead, the difference between the ways in which different lifestyles experience the process of “gestalt formation” is so great that it is impossible to conceive of the same objects as a result of processes of abstraction (or to conceive of the same subjects, one is tempted to add).

Self-realisation, which plays a central normative role for Naess and is something he would like to see supported in general, not only with respect to humans, but also in relation to non-human natural individuals, also crucially depends on gestalt

³³ Naess (1989, 95).

³⁴ Naess (1989, 45). See also Naess (1989, 163): “In the face of increasing environmental problems, the solutions proposed... revealed two trends, one in which it was presumed that a piecemeal approach within the established economic, social, and technological framework is adequate, another which is called for critical examination of the man-nature relation and basic changes which would affect every aspect of human life.”

³⁵ Naess (1989, 60 f).

formation. According to Naess, recognising oneself in other natural forms of life means that, on the one hand, something that extends beyond the limited conventional self (which Naess calls the “ego”) is realised in gestalt formation. On the other hand, this recognition brings about the actualisation of what we are as living beings, both in terms of perception and possibly also in terms of action. Naess is here inspired by Spinoza’s essentialism, according to which the individual essence of a single finite object is rooted in one infinite substance insofar as the infinite substance is the immanent cause of everything that emanates from it.³⁶ For Naess, self-realisation is therefore bound up with our understanding that we have emerged out of this totality together with other things: “So we are more than our egos, and are not fragments, hardly small and powerless. By identifying with greater wholes, we partake in the creation and maintenance of this whole. We thereby share in its greatness.”³⁷

For Naess, this partaking in a greater whole by understanding that others are like me presupposes the conception of individuals as parts of the natural-historical process through which life on earth could evolve. Understanding that each natural individual is what it is only as a result of the unfolding of a general natural-historical process should at least make it possible for humans to *identify* with those who have also taken part in this process. In a certain sense, this account can be interpreted as the historicised digest of Spinoza’s conception of the relation between finite modes and the one and only substance. In Spinoza, too, it should be possible for finite human beings to perceive the world *sub specie aeternitatis*, that is, from a perspective that captures the whole picture (or landscape) and might be identified with the point of view of the Spinozistic substance. This substance is the reason why all finite beings are possible and it can therefore be recognised in everything.³⁸

Consequently, one significant goal of Naess’ attempt to change civilisation with his deep ecology consists in the recognition of the *intrinsic value of landscapes, plants and animals*, which, according to him, are just as much a part of the greater process of life as humans are. These natural entities should therefore not be regarded as valuable only insofar as they are valuable *for human beings*, or insofar as they matter for humans in terms of their rights and obligations (which constitutes the much criticised *anthropocentric* perspective); rather these entities should be seen as independent beings and realities in which the process of life manifests itself in general: “The value of non-human life forms is independent of the usefulness these may have for narrow human purposes.”³⁹

Naess suggests that this non-anthropocentric perspective on non-human natural forms of life figures in the old mythologies of peoples with animistic religions (such

³⁶ Cf. Spinoza (1972, 18).

³⁷ Naess (1989, 173).

³⁸ See Naess (1989, 172–181) and Naess (1973, 99). Peter Rohs demonstrated that Spinoza’s substance can be interpreted in terms of field theory in: *Feld, Zeit, Ich. Entwurf einer feldtheoretischen Transzendentalphilosophie* (Rohs 1996). The relevant passages in Spinoza’s *Ethica* can be found at V, p. 24, p. 25 and p. 29.

³⁹ Naess (1989, 29).

as the Sherpa people in the Himalayas) that have presumably not yet been rendered accessible to philosophical conceptualisation.⁴⁰ In defending such an approach, Naess is not simply “re-mythologizing” our worldview; after all, even in the analytic tradition, which is far removed from any kind of mysticism, panpsychism becomes associated with ecology in a similar way as in Naess.

Especially in the context of today’s philosophy of mind in which the hypothesis of minds as having emerged from non-mental forms of life has lost much of its plausibility, we can observe a trend back to arguments defending panpsychism with the aim of establishing the claim that all physical processes involving the transmission of energy also involve sentience.⁴¹ Thus, Godehard Brüntrup has recently maintained in his presentation in Frankfurt⁴² that a panpsychist perspective of nature satisfies the normative demand of ecologists that natural forms of life be considered, not as available, but instead as *unavailable* to the satisfaction of a certain degree of human needs. It is for this reason that panpsychism would be much better suited as a theory to ground our intuitions about the unavailability of natural forms of life in an ontology than dualistic or materialistic theories of mind are. In a similar context, David Skrbina has argued that panpsychism represents a persistent undercurrent of western thought which ought to be made available to ecological argumentation, an opinion he has also tried to promote practically while running as a candidate for the American Green Party (2006) in Michigan.⁴³

At this point the following questions arise: what kind of constructs actually are “panpsychism,” “materialism” and “dualism”? How do these constructs relate to the normative concepts of everyday life? And what kind of role do they play in discourses independent of philosophy, for example in scientific, mythical or also political contexts? The legitimacy of these questions is apparent to everyone who has witnessed philosophers “declaring their commitments” to materialism, dualism, panpsychism or “real physicalism” at conferences.⁴⁴ No physicist would confess his adherence to a physical theory in this way. Thus, the fact that philosophers tend to offer such confessions apparently reveals that the broad philosophical “constructions” have a cultural function that reaches beyond academia.

To consider the function of these philosophical constructs is important, since in its attempt to change our beliefs and way of life deep ecology also makes ample use of such constructs. After all, as has been explained above, it is by way of arguing for a gestalt-theoretical panpsychist ontology that deep ecology intends to change our lifestyle and make us adopt habits which no longer embrace the belief that other

⁴⁰ Naess (1989, 61).

⁴¹ In more recent debates in the philosophy of mind David Chalmers’ *The Conscious Mind: In Search of a Fundamental Theory* (Chalmers 1997) can be seen as having triggered a new interest in panpsychism.

⁴² This talk was given at the conference of the German Whitehead Gesellschaft, 16th January 2011 in Frankfurt am Main.

⁴³ Skrbina (2005).

⁴⁴ Galen Strawson emphatically “committed” himself to “real physicalism” at a conference on panpsychism in Munich, 21st June 2011.

natural forms of life serve the mere purpose of satisfying human needs. But here one must ask if it is possible at all for philosophy to have an influence on worldviews, forms of life, lifestyles, culturally diverse evaluations or what Naess refers to as “frameworks”. And can philosophy in its cultural importance be conceptualised as an analogue to religion?

The status of panpsychisms – considered either as philosophical theories or worldviews – can be explicated by making use of the so-called man-world-relation typology introduced by Wilfrid Sellars in his essay “Philosophy and the Scientific Image of Man” from (1960/1963), as this genealogical typology suggests that there is a panpsychist phase in the unfolding of human history.

11.6 Wilfrid Sellars’ Three Images

In this paper, Sellars distinguishes between three images that man has of himself as a being in the world, namely the original, the manifest and the scientific images, while presenting these pictures in a quasi-historical chronology. As is often the case in philosophy, this chronology is not borne out by historical facts. Therefore, it is not possible to say exactly when and where one particular image prevailed or was superseded by another. Even so, Sellars’ story is not a fictitious, legitimising or delegitimizing genealogy in the sense of Hobbes’ state-of-nature/social-contract theory or Nietzsche’s genealogy of morals either. As Bernard Williams has shown, in such genealogies historical fictions are introduced for the purpose of legitimising a current political or moral practice (as in Hobbes) or to criticise it (as in Nietzsche).⁴⁵ It is evident, then, that Sellars’ chronology of images cannot be understood as a genealogy of this kind, given that Sellars believes that the *present*, that is, twentieth century North America and Europe, is governed by *two* coexisting images, namely the manifest *and* the scientific images. This coexistence is not peaceful, however, since Sellars thinks that these images are incompatible. In order to understand how this is possible, these images have to be interpreted as *systems of propositions*, which renders it notoriously difficult to understand what these images are. I will address this problem in greater detail below.

All three of Sellars’ images deal with man and the world, and Sellars’ is particularly interested in the question of whether there are differences between the three respective ways of talking about humans, on the one hand, and the non-human world, on the other. With respect to the original and the scientific image, there is said to be either no sign of such differences or that people try to avoid distinguishing the human and extra-human world in their talk. By contrast, the manifest image shows clear signs of these two fundamentally different ways of talking.

Sellars’ conception of “the original image” signifies something we would readily call *mythical animism*, since the entirety of reality is regarded as animate. People

⁴⁵ Cf. Williams (2002, ch. 2).

committed to this image believe that the wind blows where it wants to, implying that the wind is a being with an internal mental world and, indeed, a will of its own.⁴⁶ Thus, people are convinced that rivers can be sympathetic or unsympathetic to humans; that mountains can demand respect and trees care for the creatures in their surroundings, etc. As a result of these mythical beliefs, the internal world of human beings and the actions caused by them are not considered as special, plainly because every part of reality has an internal world and changes in accordance with this world. As such, there is no distinction between causality and action because the wind, water, mountains, plants and animals act as much as human beings do.

Sellars must believe that at this stage man *qua* man does not yet exist, because the original image suggests that all thought is imagistic and analogical, but not conceptual. To claim that this image skirts the anthropocentric perspective is therefore not possible because, rather than avoiding anthropocentrism, this image is based on beliefs that *precede* anthropocentrism.

Sellars conceives of the transition from the original image to the manifest image as an eminent change in category: "When primitive man ceased to think of what we called trees as persons, the change was more radical than a change in belief; it was a change in category." (Sellars 1960/1963, 10) If it is correct to claim that the mythical-animistic image is pre-conceptual, while also claiming that a change in category takes place when moving from this image to manifest thinking, Sellars' original, intuitive thinking presents itself as already containing an intuitive, categorical structure.

According to Sellars, man emerges only once he is aware of himself as a special being and able to follow rules, in particular semantic rules concerning the use of concepts. The transition from the original image to the manifest image is discontinuous because it involves that people abandon the distinction between two classes of beings: those who have obligations, who are determined by norms, and possess an internal world and act in line with their volitions and beliefs, and those constituted by objects without an internal world. According to Sellars, this discontinuity is caused by the fact that the system of conceptual norms is *holistic*. Thus, the meanings of concepts mutually refer to one another. Allegedly, a holistic system of this kind cannot develop by changing one by one but must come into existence as a whole, as if planned. It is difficult to tell whether or not Sellars already attributes a *conceptual* status to the original image (even though consistency requires him not to do so). Be that as it may, Sellars considers conceptual systems as holistically organised, while claiming that it is a condition of manhood to be in possession of concepts:

To be able to think is to be able to measure one's thoughts by standards of correctness, of relevance, of evidence. In this sense, a diversified conceptual framework is a whole which, however sketchy, is prior to its parts which are already conceptual in character. The conclusion

⁴⁶ "Thus, in the original image to say of the wind that it blew down one's house would imply that the wind either decided to do so with an end in view, and might, perhaps, have been persuaded not to do it, or that it acted thoughtlessly." (Sellars 1960/1963, 12)

is difficult to avoid that the transition from pre-conceptual patterns of behavior to conceptual thinking was a holistic one, a jump in level of awareness which is irreducibly new, a jump which was the coming into being of man.⁴⁷

Given that this argument is also used by the critics of evolutionary theory in order to challenge the claim that organisms have gradually evolved, it is hardly convincing. Be that as it may, Sellars believes that the original image must have collapsed at some point and that this has brought into existence the manifest image. This manifest image is characterised by the belief that the human world – in which persons plan, act, think, have obligations towards each other and distinguish between truth and falsehood – is very different from the world of mindless and causally responsive objects. As the reason for this collapse Sellars cites *explanatory difficulties* that essentially characterise the original image.

This consideration reveals that Sellars does in fact understand his images as *systems of propositions* used for explanatory purposes. After all, he conceives of the transition from the original to the manifest image in the following way: since the animistic view does not offer explanations of phenomena such as the regular flooding of a valley that are as compelling as explanations based on the belief in the existence of causal laws, mythical animism had to be abandoned. Since for Sellars mythical animism also performs an explanatory role, the difference between the two images is not constituted by the fact that the first image supports ritual actions, while the second furthers an interest in explanations. As we have seen, Sellars takes it that our interest in explanations is universal and manifests itself in all three images. Thus, in relation to the emergence of the scientific image, he states that here too it is the success of causal explanations that were applied to human beings that led to its establishment. If, Sellars claims, at some point, brain-physiological explanations of human actions were so successful that it made the talk about man as a norm-guided being with an inner world superfluous, the scientific image would eliminate the concept of man as much as mythical animisms did.

Even so, on the one hand, we want to keep benefitting from the explanatory success of the scientific image, and even increase this success, but, on the other hand, we also want to continue to think of humans as norm-guided creatures. This gives way to a conflict, or, as Sellars puts it, a philosophical task: the task of developing a “synoptic vision” able to unite the conception of man as norm-guided with the conception of a causally structured world.

This is not the place to discuss Sellars’ account of such a synoptic vision of this kind, since this would lead us into the technical intricacies of his “process philosophy” as developed in his *A Metaphysics of Pure Process*.⁴⁸ Instead, I will evaluate whether Sellars’ account of images (the mythical-animistic, the normative-personalistic and the scientific, causally-deterministic image) as explanatory systems of propositions is plausible. I think it is not. Regardless of this question, however, I think that in order better to understand what is at stake in Sellars it is necessary to focus more on

⁴⁷ Sellars (1960/1963, 6).

⁴⁸ Sellars (1981, 1–90).

the concept of images and to relate them to what Wittgenstein calls a “form of life” and Dilthey and Jaspers dub “worldview” (*Weltanschauung*). By identifying these worldviews (or images) with explanatory systems of propositions, it will then become immediately evident how philosophy can be used to establish worldviews. After all, reasoning, assessing arguments and explaining is the chief exercise of philosophy. Philosophy therefore seems suited to the task of examining, evaluating and criticising existing images/worldview, and can, namely in its capacity to generate great systems of thought, even be expected to create images out of its own accord.

Against the backdrop of the Marxist criticism of ideology, Dilthey’s philosophy of life and Cassirer’s philosophy of symbols, Sellars’ view of images seems naïve, because it assumes that humans have no other way of connecting with the world other than in their explanatory and assertive practices. However, such practices are not that relevant with respect to ideologies, worldviews or mythical and religious symbolic forms. Unjustified value statements and power relations also count, perhaps even more than the other elements. This consideration suggests that Sellars’ conception of languages as constituting the broader contexts of human thought and action is based on a *scientific misunderstanding*. The reason why languages can perform this role is that shared practices of human collectives manifest themselves in the establishment of broader linguistic contexts where the axioms of the relevant propositional system are taken to be self-evident. Even normative statements such as “life is sacred” can be part of such self-evident axioms. The foundational character of these axiomatic statements is here not derived from their explanatory power but instead derives from the role they play in the everyday life of people, for instance in ritual practices. If a person has been taking part in rites that stress the sacredness of life since her childhood, she will not doubt the explanatory power of the claim that life is sacred.

We have already seen that Jaspers understands such cases in the context of his theory of *evaluation* as exemplified in his *Psychologie der Weltanschauungen*. It is hierarchies of values – whatever forms these might take in a conceptual or pre-conceptual, mythical context – that guide action. And is the need for guidance of action not much more fundamental and widespread in the human world than the need for explanations, which Sellars considers as universal?

Even if we find Jasper’s language alienating, it is important to note that worldviews are more than contexts of knowledge, namely attitudes that contain value judgments, paradigmatic experiences and references to unquestioned structures of the physical world, which provide us with orientation. Nineteenth century nomads travelling through the prairie (North American Sioux, for instance), while fighting with other people embrace different hierarchies of value and experience different kinds of objects than a twenty-first century investment banker in London. Courageous fighting, hunting in changing weather conditions and the manipulation of household items are every-day activities for the Sioux, but are absent in the routines of the investment banker. Differences in linguistic spheres can thus be conceptualised as going back to differences in lifestyles.

We can here see that it is not the lack of explanatory power but a difference in evaluations and the actions resulting from these evaluations that distinguishes the Sellarsian manifest image from the mythical image. If one agrees with Ernst Cassirer's claim that the myth "exclusively holds onto the presence of its object, in the intensity with which it seizes and takes possession of consciousness in a specific moment,"⁴⁹ then it would seem that myth has nothing to do with explanation. This is because explanations require that we abstract from the particularity of the given by relating this particularity to something absent and by subsuming it under a general law. It is precisely this kind of abstraction that is necessarily involved in explanation and expresses a disregard for the particularity of the manifold to which deep ecology objects.

11.7 The End of Philosophy and the Rise of a Narrative Culture

Despite its activist elements, Naess work still depicts deep ecology as a theory, even though theory formation is no longer considered an appropriate way to capture the talk about natural individuals whose existence is significantly shaped by the experience of an inner world and who value themselves for what they are. If we want to speak about the person as such and the way she experiences her inner world, we need to *narrate* her life. In this narrative (a historical narrative, for instance), we trace her experiences and in this way find out about her. So the question is this: if deep ecology is interested in creating a form of experience that makes it possible to appreciate natural beings as values in themselves, would it not have to take recourse to narrative techniques?⁵⁰

Using examples from fiction like the works of Nabokov and other modern authors, Richard Rorty has demonstrated the importance of narration in influencing the way in which we perceive other people and the effects of our own actions on others, especially when they are *cruel*.⁵¹ Stories like *Lolita*, *Uncle Tom's Cabin* or *Les Misérables* can "help us see how social practices which we have taken for granted have made us cruel."⁵² Cruelty towards members of another race is often a consequence of the inability to adopt their perspective. Rorty does not deal with the metaphysical problems connected with the demand to cultivate the ability to adopt the perspectives of others. Nevertheless, from a *solipsistic* point of view – which does not conceptualise others as persons with a subjective internal world and leaves

⁴⁹ Cassirer (1994, 47).

⁵⁰ Films about nature can perform a similar role. I would like to thank Cornelius Brock, who pointed out to me the importance of film narratives in Kiel on 17th October 2011. For a discussion of the relevance of narratives for human life see Schapp (1953) and Lübke (1972).

⁵¹ Rorty (1989, 141–168).

⁵² Rorty (1989, 141).

open the possibility that they are zombies or machines – it is hard to understand what cruelty actually means. One cannot develop sympathy with a zombie or machine. Likewise, sympathy with the harms endured by other non-*human* forms of life seems to presuppose that we perceive them not only as objects but also as subjects, not only as an “It” but also as a “You”.

For Rorty, cruelty mainly has to do with *humiliation*. It seems to play a less crucial role in the relationship between human beings and animals, even though the denial of accepting an animal’s full range of possibilities as a living creature is conceptualised as a consequence of the “use” of vertebrates as pre-processed foodstuffs. But the claim that non-human beings are “some of us”, just like human beings with a different race are still considered in this way, can be supported by a panpsychist metaphysics (which Rorty, of course, has no intention to argue for). And conversely, a way of life which regards other non-human forms of life not as a resource but as fellow beings (as in Sellars’ “original image”) can be expected to increase the plausibility of a panpsychist metaphysics.

Now, going back to the Stoics and the Enlightenment, it would appear that there is no shortage of theories focusing on the question of what is universal to humans regardless of racial differences. Despite this, however, the impact that these theories have had on altering the way humans live and inflict pain on others is smaller than that of those social novels mentioned above. So, there is no reason for optimism when the practical impact of a panpsychist metaphysics is to be considered. Perhaps the relation between philosophy and literature is governed by principles similar to those constitutive of the relation between science and technology, considering that it is only through the effects of applied technology that scientific theories change the reality of life. Unfortunately, ever since philosophy’s Platonic beginnings, its relation to literature has never been good.⁵³ This would need to change if the engagement with narratives were to count as promising at all as a method in the philosophical endeavour of transforming our current way of life.

Maybe philosophy has had an *indirect* effect by having inspired the authors of the above-mentioned novels. It would seem that deep ecology must strive to achieve something like this, namely to put into practice an educational programme (Wittgenstein called it the ‘education of grown-ups’) with regard to the perception of non-human forms of life. But does this mean that deep ecology must take up the stories of Chinghiz Aitmatov’s ‘When the Mountains Fall’ or the books by Jane Goodall and Diane Fossey?⁵⁴

Such a philosophy would of course have to do next to nothing with ontology. If ontology is considered as the branch of philosophy that deals with existences, as it is usually done, deep ecology’s commitment to ontology emerges as striving for the recognition of the internal world of natural beings as a fact. But if this view can be developed only in an approach that combines descriptive narrative texts, on the

⁵³ Cf. Gould (1990).

⁵⁴ Or, more recently, Reichholf (2011). I have tried to connect philosophical reflection with narration related to nature in Hampe (2011).

one hand, and philosophical and scientific treatises, on the other, then the change deep ecology wants to bring about is much more fundamental than it may appear at first glance. Perhaps it is also more fundamental than its representatives themselves may think. This is because deep ecology does not target specific philosophical fields, for instance, applied ethics, that is, environmental ethics, or ontology, but instead aims for a reconciliation of science, philosophy and narration, which Rorty would describe as the attempt to turn a culture of scientific explanation into one which describes and narrates.

In his philosophy of nature, Paul Feyerabend characterises Parmenides as the beginning of the new “philosophers’ world”.⁵⁵ According to him, the worlds of Homer and the Stone Age were diverse with narrations of a range of independent creatures. By contrast, with Parmenides, the habit of thinking in abstract, non-intuitive terms was born to the effect that the importance of our concrete engagement with a multitude of beings was pushed into the background, ultimately giving way to our scientific culture. Feyerabend laments the *elimination of diversity* brought about by this culture.⁵⁶ Just as Heidegger traced back to Plato the fall of culture into a state of the “oblivion of being”, Feyerabend sees in Parmenides the beginning of the cultural fall that crucially involves the homogenising of thought, the elimination of non-intuitive explanations and the loss of the ability to appreciate the inner world of natural creatures.

But if this is the way the new thinking of philosophers and scientists can be characterised, then deep ecology attempts – in a Feyerabendian vein – to return to a perspective that appreciates the particularity of existence, and by doing so seeks to terminate the process that has begun with Parmenides and has led to the emergence of the *new world* of philosophers and scientists. This return need not be seen as a return myth, as Feyerabend stresses, as it is conceived as a turning back towards the perception of the concrete in the diversity of the manifold world.

Given the centrality of the concept of knowledge, this account will raise eyebrows: narratives, many will protest, do not convey knowledge. But this only holds true if we conceive of knowledge in a certain way, namely as something which requires a certain methodology, reproducible procedures and findings, and the applicability of general standards of justification. In short, it requires us to commit to the concept of knowledge employed in natural sciences.

But it seems that the witness who tells the judge what she has seen also knows something, even though her knowledge cannot be reproduced in experiments and her justification is amenable only to the rules of argumentation and natural language and cannot be rephrased as mathematical deductions. And is it ruled out that I *know* more about myself after having read Max Frisch’s *I’m Not Stiller*? Do I not *know* more about the lives of migratory birds once I have watched Jacques Perrin and Jacques Cluzaud’s “Winged Migration”?

⁵⁵ Feyerabend (2009, 163–171).

⁵⁶ Cf. Feyerabend (1999).

In his natural philosophy, Paul Feyerabend makes reference to the knowledge of the Polynesians, who were evidently capable of navigating long distances across the open sea without using the coastlines for orientation. In their non-literate culture, they fixed their knowledge about the use of stars for purposes of orientation in stories about appearing and vanishing gods in the sky. Stories like these were easier to memorise and pass down than abstract discourse. Later, the nautical significance of these stories was apparently lost. But according to Feyerabend, this example also shows that it is possible for narrations to encode and pass on knowledge.⁵⁷ The idea that art only delights while science instructs is simplistic. Doesn't the saying go: "fabula delectat *et* docet"?

Philosophy has always been its own critic. Perhaps it is this critical spirit that distinguishes it as a project that has its roots in Plato's "Sophistes" from other intellectual projects, such as poetry and the sciences. Since Rousseau, Nietzsche and Heidegger, philosophy's self-criticism has turned into a criticism of culture; a culture that it sometimes even believed to have created itself. Philosophy certainly can change philosophy. But if it is not also able to create culture from within itself, if this is something only religions, narratives and the applied technologies of science can do, then there is no hope that deep ecology with its ontology will be able to create something new (a theory, lifestyle, narrative culture) and transform civilisation. But this threat does not make the concerns of deep ecology implausible or incomprehensible, but still it can be efficient only if it uses narration and becomes part of a narrative culture, it cannot produce by itself. The idea that the transformation of civilisation is controllable remains philosophical megalomania. Until in our civilisation a revalorisation of a narrative culture takes place, through which our natural connectedness with the manifold diversity of particular existences can re-established, we can still rely on sensible and sometimes very effective arguments from environmental ethics which operate within the culture in which we are living today. Beyond this, one can always hope for something better.

Acknowledgments I would like to thank Isabel Adey and Naomi Osorio-Kupferblum for translating this essay from the German.

References

- Agar, D. 2001. *Life's intrinsic value*. New York: Columbia University Press.
- Aitmatov, Chinghiz. 2006. *When the mountains fall*.
- Aitmatov, Tschingis. 2007. *Der Schneeleopard*. Zurich: Unionsverlag.
- Birnbacher, Dieter. 2001. Sind wir für die Natur verantwortlich? In *Ökologie und Ethik*, ed. Dieter Birnbacher. Stuttgart: Reclam.
- Cassirer, Ernst. 1994. *Philosophie der symbolischen Formen. Zweiter Teil. Das mythische Denken*. Darmstadt: Wissenschaftliche Buchgesellschaft.

⁵⁷Feyerabend (2009, 46 f).

- Cavell, Stanley. 1979. *The claim of reason. Wittgenstein, skepticism, and tragedy*. Oxford: Oxford University Press.
- Chalmers, David. 1997. *The conscious mind: In search of a fundamental theory*. Oxford: Oxford University Press.
- Descartes, René. 1973. *Meditationes de Prima Philosophia*. In *Oeuvres de Descartes*, publiées par Charles Adam, and Paul Tannery. Vol. VII, Paris: Vrin.
- Feyerabend, Paul. 1999. In *Conquest of abundance: A tale of abstraction versus the richness of being*, ed. B. Terpstra. Chicago: University of Chicago Press.
- Feyerabend, Paul. 2009. *Naturphilosophie*. Herausgegeben und mit einem Vorwort versehen von Helmut Heit und Eric Oberheim. Frankfurt am Main: Suhrkamp.
- Fossey, Dian. 1983. *Gorillas in the mist*. New York: Houghton Mifflin.
- Fossey, Diane. 2000. *Gorillas in the mist*. Boston/New York: Houghton Mifflin Company.
- Freeman, Anthony (ed.). 2006. *Consciousness and its place in nature: Does physicalism entail panpsychism?*. Charlottesville: Imprint-Academic.
- Freud, Sigmund. 1930. *Civilization and its discontents*. London: W.W. Norton.
- Frey, Raymond. 1997. Rechte, Interessen, Wünsche und Überzeugungen. In *Naturethik. Grundtexte der gegenwärtigen tier- und ökoethischen Diskussion*, ed. Angelika Krebs. Frankfurt am Main: Suhrkamp.
- Geuss, Raymond. 2008. *Philosophy and real politics*. Princeton/Oxford: Princeton University Press.
- Goodall, Jane. 1988. *My life with the chimpanzees*. New York: Byron Preiss Visual Publications.
- Goodall, Jane. 1996. *My Life with the chimpanzees*. Revised ed. New York: Byron Preiss.
- Gould, Thomas. 1990. *The ancient quarrel between poetry and philosophy*. Princeton: Princeton University Press.
- Hampe, Michael. 2011. *Tunguska oder Das Ende der Natur*. Munich: Hanser.
- Horkheimer, Max. 2002. *Traditional and critical theory*. In *his Critical Theory*. New York: Continuum.
- Huttenlocher, Janellen, and Nora S. Newcombe. 2003. *Making space: The development of spatial representation and reasoning*. Cambridge, MA: MIT.
- James, William. 1982. *The varieties of religious experience. A study in human nature*. New York: Penguin.
- Jaspers, Karl. 1919/1960. *Psychologie der Weltanschauungen*. Berlin/Göttingen/Heidelberg: Springer.
- Konersmann, Ralf. 2006a. *Lebensform und Denkform in der frühen Neuzeit, in his Kulturelle Tatsachen*. Frankfurt am Main: Suhrkamp.
- Konersmann, Ralf. 2006b. *Kulturrkritik*. Frankfurt am Main: Suhrkamp.
- Krebs, Angelika. 1996. Ökologische Ethik I. In *Angewandte Ethik. Die Bereichsethiken und ihre theoretische Fundierung*, ed. Julian Nida-Rümelin, 346–387. Stuttgart: Kröner.
- Krebs, Angelika. 1997. *Naturethik. Grundtexte der gegenwärtigen tier- und ökoethischen Diskussion*. Frankfurt am Main: Suhrkamp.
- Landsberg, Paul Ludwig. 1933. Rassenideologie und Rassenwissenschaft. Zur neuesten Literatur über das Rassenproblem. In *Zeitschrift für Sozialforschung*. Hg. v. Max Horkheimer. Jahrgang 2 (1933), Nachdruck München 1970: DTV, S. 388–406..
- Leist, Anton. 1996. Ökologische Ethik II. In *Angewandte Ethik. Die Bereichsethiken und ihre theoretische Fundierung*, ed. Julian Nida-Rümelin, 388–456. Stuttgart: Kröner.
- Lübbe, Hermann. 1972. *Bewusstsein in Geschichten. Studien zur Phänomenologie der Subjektivität. Mach-Husserl-Schapp-Wittgenstein*. Freiburg: Rombach.
- Marietta, Don E. 1979. Knowledge and obligation in environmental ethics: A phenomenological analysis. *Environmental Ethics* 1(3): 195–207.
- Marietta, Don E. 1997. Knowledge and obligation in environmental ethics: A phenomenological analysis. In *Naturethik. Grundtexte der gegenwärtigen tier- und ökoethischen Diskussion*, ed. Angelika Krebs. Frankfurt am Main: Suhrkamp.
- Marx, Karl, and Friedrich Engels. 1953. *Die Deutsche Ideologie. Kritik der neuesten deutschen Philosophie in ihren Repräsentanten Feuerbach, B. Bauer und Stirner und des deutschen Sozialismus in seinen verschiedenen Propheten*. Berlin: Akademie Verlag.
- Marx, Karl, and Friedrich Engels. 1970. *The German ideology: Introduction to a critique of political economy*. London: International.

- Müller, Tobias, and Heinrich Watzka (Hrsg.). 2011. *Ein Universum voller 'Geiststaub'. Der Panpsychismus in der aktuellen Geist-Gehirn-Debatte*. Paderborn: Mentis.
- Naess, Arne. 1973. The shallow and the deep, long-range ecology movement. A summary. *Inquiry* 16(1–4): 95–100.
- Naess, Arne. 1989. *Ecology, community and lifestyle*. Cambridge: Cambridge University Press.
- Naess, Arne. 1997. Die Tiefenökologische Bewegung. Einige philosophische Aspekte. In *Naturethik. Grundtexte der gegenwärtigen tier- und ökoethischen Diskussion*, ed. Angelika Krebs. Frankfurt am Main: Suhrkamp.
- Ott, Konrad. 1997. IPSO FACTO. Zur ethischen Begründung normativer Implikate wissenschaftlicher Praxis. Frankfurt am Main: Suhrkamp.
- Pfordten, Dietmar vd. 1996. *Ökologische Ethik. Zur Rechtfertigung menschlichen Verhaltens gegenüber der Natur*. Reinbek bei Hamburg: Rowohlt.
- Piaget, Jean, and Bärbel Inhelder. 1956. *The child's conception of space*. London: Routledge.
- Regan, Tom. 1997. Wie man Rechte für Tiere begründet. In *Naturethik. Grundtexte der gegenwärtigen tier- und ökoethischen Diskussion*, ed. Angelika Krebs. Frankfurt am Main: Suhrkamp.
- Reichholf, Joseph. 2011. *Naturgeschichten*. Munich: Knaus.
- Rohs, Peter. 1996. *Feld, Zeit, Ich. Entwurf einer feldtheoretischen Transzendentalphilosophie*. Frankfurt am Main: Klostermann.
- Rorty, Richard. 1989. *Contingency, irony, and solidarity*. Cambridge: Cambridge University Press.
- Russell, Bertrand. 1921. *The analysis of mind*. London: Routledge.
- Schapp, Adam. 1953. In *Geschichten verstrickt. Zum Sein von Mensch und Ding*. Hamburg: Meiner.
- Sellars, Wilfrid. 1960/1963. Philosophy and the scientific image of man. In *Science, perception and reality*. London: Routledge.
- Sellars, Wilfrid. 1981. Foundation for a metaphysics of pure process: The Carus lectures of Wilfrid Sellars. *The Monist* 64(1): 1–90.
- Singer, Peter. 1979. *Practical ethics*. Cambridge: Cambridge University Press.
- Skrbina, David. 2005. *Panpsychism in the West*. Cambridge, MA: MIT.
- Spinoza, Baruch de. 1972. *Ethica Ordine Geometrico Demonstrata*. In *Spinoza opera*, vol. II, ed. Carl Gebhardt. Heidelberg: Carl Winter.
- Strawson, Peter. 1959. *Individuals. An essay in descriptive metaphysics*. New York: Routledge.
- Theobald, Werner. 2003. *Mythos Natur. Die geistigen Grundlagen der Umweltbewegung*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Wallace, David Foster. 2006. *Consider the Lobster and other essays*. New York/Boston: Abacus.
- Williams, Bernard. 2002. *Truth and truthfulness. An essay in genealogy*. Princeton/Oxford: Princeton University Press.
- Wittgenstein, Ludwig. 1969. *On certainty*. Oxford: Basil Blackwell.
- Wolf, Ursula. 1997. Haben wir moralische Verpflichtungen gegen Tiere? In *Naturethik. Grundtexte der gegenwärtigen tier- und ökoethischen Diskussion*, ed. Angelika Krebs. Frankfurt am Main: Suhrkamp.
- Wolff, Jean-Claude. 1992. *Tierethik. Neue Perspektiven für Menschen und Tiere*. Fribourg: Fischer.

Author Biographies

Stephen Gaukroger is Professor of History of Philosophy and History of Science at the University of Sydney. His current research is centered on a long-term project on the emergence and consolidation of a scientific culture in the West in the modern era. Two volumes have already appeared: *The Emergence of a Scientific Culture: Science and the Shaping of Modernity, 1210–1685* (2006) and *The Collapse of Mechanism and the Rise of Sensibility: Science and the Shaping of Modernity, 1680–1760* (2010). Work on the third volume is underway: *The Naturalisation of the Human and the Humanisation of Nature: Science and the Shaping of Modernity, 1750–1825*.

Johannes Haag is Professor of Theoretical Philosophy at the University of Potsdam. He studied at the Ludwig-Maximilians University, Munich (Habilitation 2004), and, before coming to Potsdam, taught at the Humboldt-University, Berlin. He is the author of *Der Blick nach innen. Wahrnehmung und Repräsentation* (2001) and *Erfahrung und Gegenstand. Das Verhältnis von Sinnlichkeit und Verstand* (2007) and various articles on early modern philosophy, Kant, German Idealism and contemporary topics in theoretical philosophy.

Michael Hampe holds a chair in philosophy at the ETH Zürich and has published widely on pragmatism, philosophy of science and early modern philosophy (especially Spinoza). His books *Gesetz und Distanz: Studien über die Prinzipien der Gesetzmäßigkeit in der theoretischen und praktischen Philosophie* (1996) and *Eine kleine Geschichte des Naturgesetzesbegriffs* (2007) analyze the relation between the concepts of nature and law in historical context.

Martin Kusch holds a chair in philosophy of science and epistemology at the University of Vienna. He moved to Vienna in 2009 from Cambridge where he held a chair in philosophy and sociology of science. His research focuses on social epistemology, philosophy of the social sciences, and the sociology of knowledge. He is currently completing a book on Wittgenstein and the sciences. His book publications include *Knowledge by Agreement* (2002) and *A Sceptical Guide to Meaning and Rules* (2006).

Martin Lenz is Associate Professor of Philosophy at the University of Groningen. He studied in Bochum, Budapest and Hull (Ph.D. in 2001 in Bochum) and spent his postdoctoral period in Cambridge, Tübingen and Berlin (Habilitation in 2009). He is the author of *Mentale Sätze. Wilhelm von Ockhams Thesen zur Sprachlichkeit des Denkens* (2003) and *Lockes Sprachkonzeption* (2010) and various articles on medieval and early modern philosophy.

Antonia LoLordo is Associate Professor of Philosophy at the University of Virginia. She is the author of the books *Locke's Moral Man* (2012), *Pierre Gassendi and the Birth of Early Modern Philosophy* (2007) and various articles on Descartes, Gassendi, Hume, Locke, Malebranche, and others. She received her Ph.D. from Rutgers University.

Annette Pierdziwol is a postdoctoral fellow at the Institute for Advanced Studies in the Humanities at the University of Edinburgh. She completed her Ph.D. at the University of Sydney in 2010. She has taught at the University of Sydney and the University of New South Wales since 2009 and has published on moral philosophy in contemporary continental philosophy and in the history of philosophy, particularly focusing on Enlightenment France.

Eric Schliesser is BOF Research Professor of Philosophy at the University of Ghent. He has published articles on Newton, Spinoza, Berkeley, Hume, Smith, De Grouchy, Kant, and the philosophy of economics. He is the co-editor of *Interpreting Newton: Critical Essays* (2012).

Lionel Shapiro is Associate Professor at the University of Connecticut. He has published papers in the philosophy of language, philosophical logic, and the history of early modern and twentieth century philosophy. His research interests tend to focus on the proper roles of the concept of truth and of propositional-attribute ascriptions in theorizing about the intentionality of language and thought. Related interests include deflationary views of semantic notions and the semantic paradoxes.

Anik Waldow is Senior Lecturer in the Philosophy Department at the University of Sydney. She mainly works in early modern philosophy and has published articles on the moral and cognitive function of Humean sympathy, early modern theories of personal identity, skepticism and associationist theories of thought and language. She is the author of the book *David Hume and the Problem of Other Minds* (2009).

Index

A

Agent regret, 159, 160, 166, 169,
171, 172
Appeasement, 166–170
Archetypes, 26–28, 53, 54, 58–60, 63
Artificial kinds, 88
Association, 6, 41, 44, 47, 106,
163, 204

B

Barnes, B., 95
Baumgarten, 8, 101–113
Belief, 1, 2, 5, 6, 17, 37–50, 88, 91, 93, 94,
107, 109, 124, 125, 131, 135, 154, 164,
183, 185, 187, 192, 194, 195
Bloor, D., 91
Boyd, R., 27, 29, 32, 88–92, 95
Brandom, R., 2, 3, 6, 31, 32, 39–42
Buffon, 101, 102, 109, 111, 125

C

Casuistry, 168, 171
Conatus, 6, 38, 40, 43–48
Concept, 1, 2, 4–9, 13, 17, 18, 41, 42,
47, 48, 66, 67, 68, 70–72, 74–77,
79, 80, 82–84, 87–89, 91–96, 105,
116–127, 130, 131, 159, 161, 166,
173, 181, 182, 185, 186, 188, 189,
192, 194, 195, 196, 199, 203, 204
Condillac, 101, 102, 121, 122, 147
Conventions, 2–5, 47, 128
Craig, E., 7, 87–89, 92–96
Cultivation, art of, 156, 157
Custom, 6, 47, 164

D

Demonstration, 38, 54, 104, 123
Descartes, R., 3, 4, 6, 14–24, 29–31,
33, 41, 110, 112, 115, 181,
185, 204
Diderot, D., 101–113, 116
Direction of fit, 52–55, 58–63
Discovery, 110, 117, 127–131
Dretske, F., 93
Dupré, J., 91

E

Education
negative, 140, 151
sentimental, 8, 138, 148, 152, 156
society's, 144–151
Emotions, 38, 47, 116, 130, 150, 170
Epistemology, 2–5, 7, 17, 60, 63, 87–96,
107, 137, 147, 186, 203
Essence
nominal, 23, 28, 30, 54, 59–62
real, 5, 14, 22, 25, 28–30, 54,
59–62
Ethology, 88, 91
Experience, 7–9, 28, 47–49, 53, 61, 65, 67, 69,
70, 72, 74, 75, 78, 80, 81, 82, 84, 88,
102, 103, 115, 116, 118–125, 127–132,
143, 145, 146, 149, 153, 154, 157, 158,
181, 183–190, 196, 197

F

Fodor, J., 2, 17
Foot, P., 3
Fricker, M., 87, 96

G

Gelfert, A., 87
 Genealogy of knowledge, 87–96
 Gettier, E., 93
 Goldman, A.I., 93
 Good informant, 89, 93, 95

H

Hacking, I., 89
 Haller, A., 125
 Herder, J.G., 8, 101–113, 115–132
 Hume, D., 8, 101–113, 162–164, 166,
 172, 204

I

Ideas
 clear and distinct, 17, 21, 22, 29, 31
 confused, 20–22, 30, 49
 inadequate, 14, 43, 61
 materially false, 14, 20, 24
 of modes, 29, 30, 51, 53, 55, 57–60
 of substances, 26, 27, 31, 51,
 53, 58–60
 Image-model, 67–71, 74, 75, 80,
 81, 82, 84
 Imagination, 5, 7, 15, 65–84, 106, 115–132,
 138, 147, 148, 154, 156, 157, 161,
 163, 164, 169, 170
 Impression, 68, 72–76, 78, 79, 125, 137,
 146–150, 152, 155, 156, 157
 Induction, 27, 28
 Intentionality, 5–7, 13–33, 67, 204
 Intuition, 66, 67, 69–84, 88, 92, 94, 126, 184,
 185, 190, 192

K

Kant, I., 175
 Kappel, K., 87, 95
 Kelp, C., 87
 Kornblith, H., 7, 28, 29, 87–95

L

Language, 1, 2, 7, 17, 32, 42, 44, 54, 71, 93,
 110–112, 117–122, 124, 128, 129, 130,
 136, 145, 162–164, 170–174, 183, 187,
 196, 199, 204
 Laws of nature, 40, 174
 Leibniz, 51, 61, 62, 109, 115
 Locke, J., 4–7, 14, 15, 22–31, 33, 51–63, 101,
 103, 112, 120, 139, 147, 204

M

McDowell, J., 2, 3, 67, 68, 70, 75, 76
 Metaphysics, 2, 8, 60, 101–113, 115, 116, 120,
 122, 123, 126, 127, 131, 132, 184–186,
 195, 198
 Millikan, R., 2, 27, 32
 Misrepresentation, 2
 Mitova, V., 96
 Modes, 6, 29, 30, 38, 51–63, 107, 148, 154,
 163, 184, 191
 Moral sensibility, 108, 112, 135–158
 distortions of, 144
 Müller-Lyer illusion, 49

N

Naturalism, 2, 112, 125
 Naturalistic fallacy, 3
 Naturalized epistemology, 87–96
 Natural kinds, 6, 27, 28, 29, 52, 58–60, 63,
 88–92, 95
 Natural law, 4
 Natural normativity, 38, 40, 41, 47, 48, 50
 Natural sociality, 140, 141, 143, 144, 145,
 149, 151
 Nature, 1–9, 13, 14, 17, 20, 21, 22, 24, 25, 29,
 30, 37–40, 43, 44, 48, 49, 53, 59, 60,
 63, 65, 67, 68, 77–79, 84, 87, 88,
 93–95, 103, 104, 107, 115, 116, 118,
 121, 122, 124, 125, 131, 135–158,
 160–165, 169, 172–175, 181–184, 186,
 187, 190, 192, 193, 197–199, 203
 Neta, R., 87
 Normativism, 2
 Normativity,
 natural, 38, 40, 41, 47, 48, 50
 social, 37
 Novelty, 13, 117, 129–131
 Nozick, R., 93

O

Observation, 53, 72, 75, 115, 116, 118,
 123–126, 132, 142, 145, 147, 153, 155,
 161, 162, 163, 165, 168
 Oedipus, 167

P

Piacular, 9, 159–176
 Pritchard, D., 87
 Promiscuous realism, 88, 89, 91
 Propositional attitudes, 6, 31, 38, 41, 43, 44
 Putnam, H., 1

R

Radford, C., 93
 Rationalism, 15, 29–31, 103
 Rationality, 147, 148, 160, 162–165
 Relations, 1–9, 16, 17, 39, 41, 43, 46, 48, 49,
 51, 52, 55–59, 62, 69, 70, 73, 74, 82,
 84, 93, 101–103, 106, 108, 109, 112,
 115, 116, 117, 121, 125, 127, 130, 131,
 135–139, 143, 146, 148, 150–153, 156,
 157, 160–165, 174, 188, 190, 191, 193,
 195, 196, 198, 203
 Representationalism, 31, 32
 Rorty, R., 3, 4, 137, 139, 147, 148, 197,
 198, 199
 Rosenberg, J., 67, 70, 75
 Roszkowski, A., 96
 Rousseau, J.-J., 8, 101, 106, 107, 112, 122,
 124, 135–158, 163, 186, 200

S

Scepticism, 89, 94, 181, 182
 Schema, 68, 70, 74, 75, 112
 Second nature, 4
 Sellars, W., 65, 67–71, 75, 79–82, 193–198
 Sensation, 3, 6, 7, 19–21, 30, 52, 68, 76,
 78–84, 101, 102, 105, 106, 108, 110,
 117–122, 127, 130, 147

Sensationalism, 146

Sensibility, 7, 8, 65, 66, 71, 75, 82–84,
 101–113, 115–132, 135–158, 203
 Smith, A., 9, 95, 159–176, 204
 Social normativity, 37
 Spinoza, B., 4–6, 37–50, 125, 179, 180, 185,
 187, 190–193, 203, 204
 State of nature, 87, 88, 93–95, 139, 140,
 143, 193
 Substances, 6, 25–27, 51–63, 110
 Sympathy/sympathetic process, 103, 160–162,
 165, 170, 174, 175, 198, 204
 Synopsis, 73, 76–83
 Synthesis, 66, 67, 71–79, 81–84, 87, 120, 127

T

Testimony, 94, 136

U

Unger, P., 93

W

Williams, B., 38, 47, 87, 159, 166, 193
 Wittgenstein, L., 2, 93, 94, 189, 196, 198
 Wolff, R.P., 83