

A COMPANION TO NATURALISM

Edited by
JULIANO DO CARMO

DISSERTATIO
FILOSOFIA

A COMPANION TO NATURALISM

DISSERTATIO'S SERIES OF PHILOSOPHY

A COMPANION TO NATURALISM

EDITED BY
JULIANO S. DO CARMO

DISSERTATIO
FILOSOFIA

PELOTAS, BRAZIL

FIRST PUBLISHED 2015
BY NEPFIL
FEDERAL UNIVERSITY OF PELOTAS, BRAZIL

All rights reserved. No part of this book may be reprinted or reproduced or utilized in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including any information storage or retrieval system, without permission in writing from the publishers.



UFPEL



BRAZILIAN LIBRARY CATALOGUING IN PUBLICATION DATA

A companion to naturalism
edited by Juliano S. do Carmo
Pelotas – RS: NEPFIL online, 2016
280p. – (Dissertatio's Series of Philosophy)
<http://nepfil.ufpel.edu.br/dissertatio/indez.php>

ISBN: 978-85-67332-34-5

1. Naturalism. 2. Philosophy 3. Mind
I. Juliano S. do Carmo (Organizer) II. Series

COD 100
146
CIP

Content

Preface	7
Notes on Contributors	8
1. Responsibility Without Freedom <i>Adriano Naves de Brito</i>	10
2. Liberal Naturalism, Ordinary Things and the Importance of Second-Personal Space <i>David Macarthur</i>	19
3. The Poverty of Analysis <i>David Papineau</i>	30
4. Demystifying the Nature of Moral Judgments by Reductionist Naturalism <i>Evandro Barbosa</i>	51
5. Is a Phenomenological Perspective on the Social Construction of Reality Still Viable? <i>Fabrcio Pontin</i>	68
6. Value and Moral Pluralism <i>Flávia Carvalho Chagas</i>	91
7. Implications and Possibilities of Naturalism in Nietzsche's Philosophy <i>Felipe Szyszka Karasek</i>	101
8. The Naturalization of Morals and the Farewell to Naturalism in Nietzsche's Philosophy <i>Clademir Luís Araldi</i>	111
9. Is Ontic Structural Realism a Viable Form of Naturalized Metaphysics? Naturalised Metaphysics and Troubles Caused by the History of Science <i>Jack Ritchie</i>	121
10. Thales of Miletus and the Greek Naturalism <i>João Hobuss</i>	130
11. The Natural Kind Status of Delusion <i>José Eduardo Porcher</i>	137
12. The Role of Ostensive Training in Natural Language Acquisition <i>Juliano do Carmo</i>	161

13. On the Probability of Theism <i>Luis Rosa</i>	174
14. Naturalism and Realism <i>Mario De Caro</i>	182
15. Revisiting the Mind-Brain Pickle: Swinburne's Event Dualism of Substances and Properties <i>Nythamar de Oliveira</i>	196
16. Naturalizing Intentionality <i>Ruth Garrett Millikan</i>	211
17. Mental Elements as Part of a Scientific Architecture <i>Sofia Inês Albornoz Stein</i>	217
18. Some Metaethical Problems for Pragmatic Naturalism <i>Susana Nuccetelli</i>	229
19. Naturalism in Philosophy of Chemistry or: Why does Metaphysics of Nature Matter? <i>Thomas Sukopp</i>	238
Bibliographical References	255

PREFACE

Offering an engaging and accessible portrait of the current state of the field, *A Companion to Naturalism* shows students how to think about the relation between Philosophy and Science, and why it is both essential and fascinating to do so. All the authors in this collection reconsider the core questions in Philosophical Naturalism in light of the challenges raised in Contemporary Philosophy. They explore how philosophical questions are connected to vigorous current debates - including complex questions about metaphysics, semantics, religion, intentionality, pragmatism, reductionism, ontology, metaethics, mind, science, belief and delusion, among others - showing how these issues, and philosopher's attempts to answer them, matter in the Philosophy. In this sense, this collection is also compelling and illuminating reading for philosophers, philosophy students, and anyone interested in Naturalism and their place in current discussions.

There are no formal divisions in this Companion, so the different essays are freestanding and can be read in isolation from the other. Nevertheless, the reader can view an informal line uniting the essays, with one topic naturally leading to the other.

The editor thanks all the contributors to this volume, and makes special acknowledgment to the National Council to Scientific and Technological Development (CNPq) for financing this project. Thanks are offered also to Aristotelian Society for permission to reprint the text "Poverty Analysis" (by Prof. David Papineau).

Juliano Santos do Carmo (Editor) is Professor of Philosophy at Federal University of Pelotas. He is the idealizer of the *Naturalism: Contemporary Perspectives* (2013) and author or editor of various studies on Wittgenstein's Philosophy, Naturalized Semantics and Normativity..

NOTES ON CONTRIBUTORS

ADRIANO NAVES DE BRITO is Associate Professor of Philosophy at the University of Vale do Rio dos Sinos – Brazil. He has written on many topics concerning Philosophical Naturalism, mainly “action and value in moral naturalized.”

CLADEMIR LUÍS ARALDI is Associate Professor of Philosophy at the Federal University of Pelotas – Brazil. He has written on Nietzsche’s Philosophy and the Naturalization of Morals.

DAVID MACARTHUR is Associate Professor of Philosophy at the University of Sydney – Australia. His research interests are Contemporary Pragmatism, Liberal Naturalism, Skepticism, Metaphysical Quietism, Philosophy of Language, Wittgenstein and Philosophy of Art.

DAVID PAPINEAU is Professor at King’s College of London and at City University of New York. He has written on many topics concerning Naturalism, among them Consciousness, Rationality, Evolution, Reality and Representation. He is one of the main precursors of Philosophical Naturalism.

EVANDRO BARBOSA is Associate Professor of Philosophy at the Federal University of Pelotas – Brazil. His research interests include Metaethics and Political Philosophy.

FABRÍCIO PONTIN is a Postdoctoral Research Fellow at the Pontifical Catholic University of Rio Grande do Sul – Brazil. His areas of expertise are Political Philosophy, Phenomenology, Theory of Reasoned Action and Theory of Law. His areas of interest include Philosophy of Language, Queer Theory and Philosophy of Mind.

FLÁVIA CARVALHO CHAGAS is Associate Professor of Philosophy at the Federal University of Pelotas – Brazil. She has written on topics in Ethics, mainly the Justification of Moral Values.

FELIPE SZYSZKA KARASEK is Professor of Philosophy at the Cultural Development Institute – Brazil. His main work concerns Nietzsche’s Philosophy and the Naturalization of Morals.

JOÃO DO NASCIMENTO HOBUSS is Associate Professor of Philosophy at the Federal University of Pelotas – Brazil. He has great contributions in the field of Ancient Philosophy, especially on the “Practical Philosophy of Aristotle”.

JOSÉ EDUARDO PORCHER is a Postdoctoral Research Fellow at the Federal University of Paraná – Brazil. He has been researching about the possible learnings and contributions of philosophers regarding the study of Abnormal Cognition and Psychopathology.

JULIANO SANTOS DO CARMO is Associate Professor of Philosophy at the Federal University of Pelotas – Brazil. He has written on various topics concerning Naturalized Semantics, mainly “Natural Language Acquisition.”

LUIS ROSA is a Humboldt Research Fellow at the Munich Center for Mathematical Philosophy – Germany. His main work deals with Epistemology and Psychology of Reasoning, Epistemic Logics and Formal Semantics.

MARIO DE CARO is Professor of Philosophy at the Università Roma Tre and Tufts University – Italy and Australia. His main work is in Realism and Naturalism discussion.

RUTH GARRETT MILLIKAN is Associate Professor of Philosophy at the University of Connecticut – USA. Her research interests span many topics in the Philosophy of Biology, Philosophy of Mind, Philosophy of Language, and Ontology. Her main work is in applications of Evolutionary Theory in these areas and in the ontology that makes the study of biological entities possible.

SOFIA STEIN is Associate Professor of Philosophy at the University of Vale do Rio dos Sinos – Brazil. Her main areas of research are Philosophy of Language, Epistemology and Cognitive Sciences. Her current work concerns Philosophy and Cognitive Sciences.

SUSANA NUCCETELLI is Professor of Philosophy at St. Cloud University – USA. Her main work is in the fields of Philosophy of Mind and Philosophy of Language. She has made significant contributions to Pragmatic Naturalism.

THOMAS SUKOPP is Associate Professor of Philosophy at the University of Siegen – Germany. He has written several papers and books about the Philosophy of Chemistry, Contemporary Epistemology and Ontological Naturalism.

RESPONSIBILITY WITHOUT FREEDOM

ADRIANO NAVES DE BRITO

INTRODUCTION

Relations between philosophy and sciences have been reworked. In the 20th century these relations were weakened, be it due to scientific specialization, be it because of the humanist vocation of philosophy which inspires it to take on a normativist bias, thus aimed more at prescription than at description. After, at the end of the 19th century and beginning of the last century, physics and mathematics had a great influence on philosophy, biology, neurosciences and psychology have more recently reintroduced sciences into the philosophical reflections, with consequences that are still only timidly taken up by academic philosophy. Especially with regard to the concept of personhood, the influence of those disciplines shall be disturbing, once the biological roots of the cultural nature of humans were deeper scrutinized.

This text appreciates this new rapprochement and, in the precise sense that it espouses the idea that all phenomena of nature, be they physical or mental, can be explained only by natural forces, immanent to nature itself, it is naturalistic. It is also naturalistic in the sense that it does not deny the possibility that there are supernatural forces which determine the world. The denial of this possibility, as well as its affirmation, would go beyond the knowledge we can have of the world and would, therefore, be a transgression of the limits of naturalism. That possibility or its denial, however, do not add any explanatory value to the phenomena. Thus, they have no epistemic value and leave a frankly naturalized ontology untouched. In this sense, the naturalism that I advocate is ontological and epistemic, but not metaphysical, which means I am not concerned with entities beyond the limits of physics. From that metaphysical (i.e. transcendent) standpoint, the naturalism that I espouse has nothing to say. In the meaning that the term “critique” has for Kant’s mature, transcendental philosophy (1781), the naturalism espoused here may well be called “critical naturalism”. It is a naturalism that reaches until where possible experience reaches. Experience, it should be said, such as the sciences can handle.

It is from the viewpoint of this naturalistic approach that I shall discuss ethics and personhood, especially the relationship between responsibility and freedom regarding the ethical dimension of personhood. My purpose is to defend the position that claims that we can get rid of freedom, understood in the strong Kantian sense, without jeopardizing the capacity of persons to bear responsibilities and, therefore, without jeopardizing personhood. That position

might be considered compatibilistic, were it not for the fact that it denies freedom as a causality and, thus, voids the need to render it compatible with the forces of nature. For the same reason, this position cannot be classified within the scope of the incompatibilistic positions, since, if there is no causality by freedom, there is then nothing that might be incompatible with the natural determinations. Given that I begin with the denial of causality by freedom, which eliminates part of the problem to which I want to dedicate myself in this paper, its core task is to explain the second element of the problem, i.e. responsibility, without, however, taking recourse to freedom.

DISPENSING WITH FREEDOM

Before dispensing with freedom, it is useful to provide a better explanation of my *ab ovo* denial of it, so that the notion of responsibility which I have to account for will also become clearer. As regards freedom, I am considering positions that are well illustrated by those defended by Chisholm (1966) and Strawson (1994) and that sustain themselves due to the same background concept of metaphysics, namely the idea that responsibility implies the agent's full autonomy, ergo, the agent's capacity, in Kant's terms, to start a causal chain by themselves.

Let us first take Chisholm's view. He says the following about the conditioning factors of responsibility:

“If we are responsible ... then we have a prerogative which some would attribute only to God: each of us, when we act, is a prime mover unmoved. In doing what we do, we cause certain things to happen, and nothing—or no one—causes us to cause these events to happen” (Chisholm, 1966).

In order to establish a perspective for this position, it is worth noting that in the philosophy of Kant, but also in that of Descartes, it is precisely the property of our being free, in a metaphysical sense, that brings us closer to God. The theory of error proposed by Descartes (1641), for instance, is based on the assumption that our will is infinite, like that of God, but that our understanding is limited. The same, *mutatis mutandis*, applies to Kant (1788). As legislators of a Kingdom of Ends, we are like deities; but as knowing beings we are limited by the conditions of possibility of experience. Chisholm is, therefore, well inserted into a long line of thinkers who, in a consistent manner, think about freedom as an uncaused causality. I say consistent because only this strong, metaphysically charged notion solves the problem of responsibility without further gaps or difficulties. If persons are free in this divine sense, persons are responsible, and nothing else needs to be added to the topic. More recently Galen Strawson continued this metaphysical tradition with an argument that, although coming from the same metaphysical foundations as regards the relationship between freedom and responsibility, diverges from his predecessors regarding divinity, of human being or of any other being, and consequently derives a skeptical position regarding responsibility. He argues that

- (i) Nothing can be *causa sui*—nothing can be the cause of itself
- (ii) In order to be truly morally responsible for one's actions, one would have to be *causa sui*, at least in certain crucial mental respects.
- (iii) Therefore nothing can be truly morally responsible (Strawson, 1986).

Strawson rules out the possibility that there may be some being that is, in a metaphysically relevant sense, *causa sui* and takes a position that coincides with metaphysical naturalism, namely, that it is impossible that there are supernatural forces that determine nature. On that basis, and on the basis of a notion of responsibility that coincides with the tradition that I highlighted above, nothing or nobody can be responsible. Thus, the skeptical position is likewise consistent with the notion of responsibility that it espouses and that, as I said, is the same as the tradition for which Kant is the paradigmatic representative for contemporary philosophers. Strawson's position appears to add a further advantage thanks to the metaphysical naturalism that it incorporates, since it can play the same card as its respectable predecessors about responsibility, without, however, offending the scientific spirit of our times. In my view, this appearance is misleading. Strawson already offended the scientific spirit when he took on a metaphysical naturalism. I shall return to this point below.

I disentangle myself from these positions without undoing the metaphysical knots with which they were woven. The type of naturalism that I espouse allows me to do this, without much trouble, although this move will charge its price not in the field of freedom, but in that of responsibility. Let me explain. If the conditions to attribute responsibility imply assertions on the metaphysical status of our species as peopled by individuals who are between gods and mortals, or else assertions about divine entities themselves, these conditions cannot be met by science or by any discipline that operates within the limits of immanence. The type of naturalism that I adopted, which is epistemic in character and committed to an immanent ontology, limits the scope of the investigation and the treatment of the problem, so that it forces me to dispense with freedom. The same critical naturalism leads us to reject, precisely for the same reasons, Strawson's position which, as I made quite clear, must be committed to a statement of a metaphysical nature, even if naturalistic, but that goes beyond the limits of what the sciences can contribute, viz. that nothing can be *causa sui*. Now, no science can claim that there cannot be something that is *causa sui*.

A direct consequence of the conjunction of a critical naturalism with the arguments of this line of thinkers would be skepticism regarding responsibility. The argument would be as follows: if we cannot decide whether there are or not beings that can be *causa sui*, or whether we ourselves, at least regarding morality, cause our own actions, then we can also not decide whether someone can or not be considered responsible for their actions. As I warned, the bill for the position I take is charged in the field of responsibility. How can one respond to this? There is no response if the terms of the problem are not changed. The skeptical conclusion is inevitable if the notion of responsibility is maintained just as the tradition—that I will call the tradition of Kant—conceives and accepts it. Do we have good reasons to maintain it? My answer is that we do not.

Since the notion of responsibility with which the tradition of Kant operates depends on another notion that is equally metaphysically charged, namely that of a transcendent freedom, and which we have to dispense with from the explanation of the moral phenomenon in consonance with critical naturalism, the notion of responsibility must be also affected by this move of disentanglement. As regards responsibility, however, the point is not getting rid of it. On the contrary of freedom—taken as a *sui generis* causality and the denial of which has no consequence whatsoever for the explanation of innerworldly phenomena, which is a statement that is perfectly in consonance with epistemic naturalism—the denial of responsibility unavoidably limits the explanation of morality. Without the notion of responsibility, morality makes no sense. And this is confirmed by human moral practices.

Skepticism, as regards responsibility, may even interdict the metaphysical certainty regarding guilt, and which would require the assimilation of freedom, but it cannot limit the game of mutual evaluations that we humans play daily. Without the act of holding a person responsible, the dynamics of morality simply could not start to move. Consequently, even if in a non-metaphysical sense, we attribute responsibilities to ourselves (and even freedom in a non-metaphysical sense, but I shall return to this point only at the end). Responsibility, then, must be able to circulate in this world without its metaphysical counterpart, viz. freedom as causality, and a philosophical position that begins with this finding must offer a notion of responsibility, that, although it does not overload it metaphysically, cannot take away the role it plays in the moral phenomenon, i.e. its role as a foundation for the attribution of guilt. Moral skepticism may even motivate philosophers, but does not really affect ordinary life.

MORALITY AND EVOLUTION

Seen from an evolutionary perspective, morality is related to facilitating cooperation. It is, thus, a functional trait of collective human behavior that evolved from the behavior of the species that preceded us on the evolutionary line. Given its importance for the reciprocal control of the behavior of individuals in a group, and given the importance of group life for the *homo sapiens sapiens* species, morality is not only functional to it, but also gave it an advantage in dealing with natural selection.

The control of the behavior of individuals for a sustainable life in a group implies both the selection of preferences appropriate to this kind of sociability—as, for instance, the disposition to positive and negative affective reactions, such as indignation, guilt and moral approval vis-à-vis, respectively, the deviation of behavior or behavioral appropriateness—and the selection of neuronal structures that can correct individual behaviors. In a species with such a flexible menu of behaviors as is the human one, the mechanisms of adjustment of individuals' behavior are decisive to ensure, at the same time, adaptability and functionality of the groups (Brito, 2014). Recent studies show that the neuronal structures are influenced by the threat of punishment in a process that reinforces the susceptibility of individuals to these threats and in this way favors the institution

of rules of conduct, whose result in turn would be an increment in the capacity of cooperation among groups. These studies illuminate the functionality of the modern systems of justice based on very primitive, but spectacularly efficient evolutionary processes. The study by J. W. Buckholtz and R. Marois (2012), “The roots of modern justice: cognitive and neuronal foundations of social norms and their enforcement”, goes in this direction.¹ According to its authors, the bias created by the threat of punishment influences the reward-based decision-making mechanisms, so that the individuals would, because of this, have a greater propensity to behaviors appropriate to the rules prevailing in the group, which would make cooperation easier. This conclusion is by the authors of the study, who formulate it as follows:

The research is consistent with the hypothesis that the evolution of norm-based decision making mechanisms may have been facilitated by the presence of neural circuitry that supports domain-general cognitive processes for value-based action selection. Such processes may have been co-opted and expanded to operate in the social domain where they promote action selection according to higher-order action values linked to social rewards, such as reputation and trust, thereby facilitating cooperation (Buckholtz & Marois, 2012, p. 655).

The relationship between neuronal structures and biases and decision-making is clear when human being is considered against the background of their biological condition.² Insofar as we are animals, we are a body and a brain that make up the same whole that acts in a given direction, although it is not exhaustively determined. The direction is determined by the basic preferences that result from the traits that were reinforced in the natural selection process, because they allowed a better adaptation to the diversities of the evolutionary environment. One of them is sociability. The behavioral bias that is thus outlined, however, leaves a vast elbow room, in which the groups can move and organize their behavior in order to carry out the main tasks and challenges of an individual’s life, i.e. surviving and procreating.

In the case of genus homo, thanks to its magnificent brain capacity, these tasks and challenges were broken down into a great number of other activities and could be efficiently performed and overcome in such varied ways that the elbow room in the most successful species, ours, broadened to the point of appearing to be free from the determinations of its animal nature. Despite the

¹ “It is possible that the threat of punishment may increase cooperation by biasing reward-related action selection mechanisms mediated by dorsal frontostriatal circuitry. In this context, punishment threat may change the reinforcement contingencies associated with potential responses, and increased corticostriatal fMRI signal in the punishment condition may reflect this updating process” (Buckholtz & Marois, 2012, p. 655.)

² For a recent discussion on the repercussions of neurosciences on law, see J. Greene & J. Cohen, 2004.

elbow room that our species has to organize its life as a group, we are animals, animals which end up having a personhood in a given culture.³

As animals we are susceptible to the causal determinations that influence all living beings, but also to those determinations that are specific to us because of the type of behavior that we have to prefer in consonance with the traits that natural selection has reinforced in us. Everything according to a causal continuum that knows no exception, that is, without the need to consider any supernatural force to explain this process adequately. As Dennett explains: “Our brains have been designed by natural selection, and all the products of our brains have likewise been designed, on a much swifter timescale, by physical processes in which no exemption from causality can be discerned” (Dennett, 2003, p. 305). It is in this framework of causal determinations, in which we have our peculiarities as a species but no transcendent privilege, that responsibility must fit. And it does fit.

Two conditions are required for it to make sense to attribute responsibility to an individual who is, as Aristotle (1997) defined it, a ζῷον πολιτικόν, an animal with a social life, an animal adapted to urban life, in which interaction and cooperation among individuals is the rule. The conditions are as follows: (1) that they can act in another way, or, as we say about an action that has already been performed, that they could have acted differently from what they did, and (2) that they are susceptible to the influence of external moral pressure on their behavior. The concept of human being that I outlined above and these two conditions provide the core for a theory of responsibility that, within the limits of critical naturalism espoused here, dispenses with freedom as causality. Once freedom as causality has been dispensed with, it is not ruled out that, in the terms of Dennett (1994 and 2003), there is elbow room for different courses of action to take place, and that carrying one of them out implies an individual taken as a causal link of events, for whose consequences they can be held responsible given the conditions listed above. In the rest of this text, I explain these two conditions and show why they are compatible with what the judicial systems consider indispensable for imputability, just as they are compatible with the conditions that we take as indispensable to attribute responsibility to persons within the scope of morality.

RESPONSIBILITY AND CONTROL OF BEHAVIOR

The first condition to render an individual responsible is that which distinguishes a voluntary action from an involuntary one. The matter is not merely physical or biological—that is, whether the individual was causally implicated in the course of ongoing actions and had biological control of their actions—but also psychological in a subjective sense, that is, in the sense that the individual could exert influence, by their capacity to make choices and at the limits of a reasonable pressure, on the direction of the facts that conditioned their action. If the physical and biological conditioning factors imposed the action on the

³ Humans are animals in the precise sense proposed by the doctrine of animalism according to its best known advocate. Eric T. Olson says: “When I say that we are animals, I mean that each of us is numerically identical with an animal. There is a certain human organism and that organism is you. You and it are one and the same” (Olson, 2003, p. 1).

individual, or if the psychological pressures were so great that it would not be reasonable to expect individuals generally to resist them and act in another way, then the individual is non-imputable, meaning that it does not make sense to require from them that they should have acted in another way. If the limiting conditions that have just been presented are not given, then the individual is imputable. The reasoning, once the due adjustments have been made, can be applied to omission. An involuntary omission does not imply imputability, whereas a voluntary omission does. Voluntary, therefore, is not defined through a causality by freedom, but through natural causality. The development of facts required that an individual take action in the sense that they were a link in the causal chain and had elbow room available to choose more than one course of action, according to the forces that preponderated in it.

The second condition is connected to the first as a specification of the type of requirement which pertains to responsibility, i.e. a moral requirement in the broad sense, that is, which also includes the legal requirements. If the individual could act in another way (also as regards omitting themselves or not) and if they were susceptible to moral pressures, in other words, if the prevailing normative requirements could preponderate in their decision, then they are responsible for their acts, and it makes sense to consider them responsible precisely because it made sense to demand of them that they should have acted in a specific manner and under the then prevailing moral—in a broad sense—constraints. Under these conditions, responsibility, as well as the will as I described it previously, do not demand causality by freedom but, conversely, they assume that the individual is willing to be determined. In other words, the disposition to have behavior guided by the moral pressure that the socially or legally constituted group exerts on them. Holding a person responsible is itself a fundamental part of the pressure that is exerted on the individual. It is because they anticipate its possibility that they guide their decisions appropriately, in a mechanism that may well be explained by the same neuronal structures and processes pointed out in the study by Buckholtz and Marois (2012) that I cited earlier. Paraphrasing another Strawson (1962), Peter, the practice of holding a person responsible does not only exploit our nature, it expresses it.⁴

In brief, the responsibility of an individual depends on their action—understood in a broad sense so as to include omission—being voluntary in a sense that does not imply freedom as causality, but takes into account the causalities involved in the circumstances of the action. An individual deliberated under the effect of the causalities which have an influence over their action, with or without an awareness of all of them, and the judgment as to whether their action can or not be considered their responsibility takes into account the force of these influences and not the individual's capacity to start, without the concurrence of those influences, a causal chain through an act of freedom. Once the conditions are evaluated by those who judge the case—morally or legally—

⁴ “What is wrong is to forget that these practices, and their reception, the reactions to them, really are expressions of our moral attitudes and not merely devices we calculatingly employ for regulative purposes. Our practices do not merely exploit our natures, they express them” (Strawson, 1962, p. 70).

if the action is considered involuntary, then the responsibility of the individual agent will be excluded. This evaluation is subjective, since it takes into account the agent's intention; but it is also objective, and considers the facts and constraints that influenced or should have influenced the agent, including the very pressure of the then relevant and prevailing normative requirements. The objectivity of the evaluation highlights precisely the second condition to hold the agent responsible, i.e. that they have the disposition to be affected by the external requirements of other individuals or of the system of norms to which they are submitted. The idea that the agent should have been influenced differently by the circumstances is essential for holding them responsible, but the very act of holding responsible, or its possibility, is considered a factor that should have influenced and finally determined their action. In this way, even if the agent has a second order wish which is identical in content to the first order wish to commit an illegal or objectionable act, if the circumstances and her dispositions are such that she could have been influenced, concerning her second order desire too, differently by the circumstances, then it makes sense to consider her responsible. In this case, the agent's bias to desire the reprobable desire is not an excuse for her not to be imputed. The point is that she should have been influenced differently also in her second order desire by social and legal norms. Here the burden of a mature personhood takes its price. This conclusion agrees with that of Frankfurt (1971), although it comes from a different meta-ethical position.

Now, once the second condition exists, namely the disposition to be effectively affected by moral pressure in the broad sense, it is appropriate to evaluate the merit or demerit of the individual's action in order to find to what extent imputability can be attributed to them; in other words, it should be determined to what extent their disposition is susceptible to pressures for a different behavior than the one they chose. Here one must find out whether there was deceitfulness or malice. If deceitfulness or malice cannot be found, the objective conditions may be suspended, since it is no longer possible to determine whether the agent could react as the current conditions, norms and pressures required them to act, which includes the possibility that they might be considered responsible for the result of the action. The mature capacity for deceitfulness and malice is also a sign of a mature personhood.

I resort to paradigmatic cases of the legal system to explain this point. For instance, let us look at the case of juveniles. Although, when considering a given reprehensible result, one might suppose that they could, under those circumstances, have acted in another way, imputability is impaired by the uncertainty regarding the efficiency of social and moral pressure on them. The matter of deceitfulness or malice cannot be investigated, because the agent cannot yet be considered someone who can conjugate, as expected, the circumstances and normative requirements prevailing at the time of the action. Under these conditions, the evaluation of deceitfulness or malice is suspended and, therefore, also the evaluation of guilt. Considering the stage of neurological development of juveniles, the best explanation to diminish their responsibility is not the fact that they cannot freely start causal chains, but the fact that they are

still poorly guided under moral pressure. This means that the moral pressure exerted by the social group on the juveniles can result in behaviors different from expected, without it being possible to determine the guilt of individuals as regards this result, so that investigating guilt also loses its object, i.e. an objective ensemble of circumstances and an agent who is predictably susceptible to normative pressures. The same reasoning can be applied to another paradigmatic case, viz. psychopaths. Even though they can understand the rules, their susceptibility to the moral pressures of the group is not enough to predict a behavior according to these pressures, with an acceptable guarantee for the group that can be harmed by the psychopath. They might be capable of deceitfulness or malice, but in an artificial way, since they cannot feel (and therefore fully understand) the burden of guilt that should accompany those who have developed a mature personhood.

The case of Phineas Gage, made famous by Antonio Damásio (1994), is a clear illustration of this point. Gage's behavior was affected by an accident that destroyed a large part of his left frontal lobe, and he was no longer responsible for his acts, not because he did not cause them, but because his susceptibility to the moral pressures of society was affected so that his behavior became unpredictable and could no longer be determined by the pressure exerted on him by the others. Because of this, even the act of holding him responsible was no longer effective as a behavioral constraint. Gage's mind could no longer be influenced towards the group's expectations, but was influenced by forces that led his behavior to disagree with those expectations and disagree in a way that was unpredictable or could not be corrected by ordinary moral pressures, including that of responsibility. Moreover, he lacked the capacity to feel guilt for not acting in accordance with the social influences he was exposed to. His personhood was impaired.

IN CONCLUSION

The objective of this text was to present the outline of a naturalized view of the notion of responsibility, such that the latter could be understood in terms of natural and immanent causality, that is, without the assumption of a causality by freedom. From the point of view of a critical naturalism, as I defined it, responsibility can be explained based on the disposition of the agent to be appropriately influenced, that is, influenced according to the mutual demands that can be made from them by members of society and that in turn are guided by reasonable moral principles. That reasonability, considering also the naturalist perspective, is aligned with the principles and preferences that favor cooperation and that constituted an evolutionary advantage for our species. It makes sense to hold individuals responsible if they are susceptible to the social influences, amongst which precisely is the act of holding a person responsible, whose anticipation should, by means of guilt and as expected by the other members of the social group, influence the decisions of an individual towards aligning them with the prevailing norms, whether contained in statutes or not.

LIBERAL NATURALISM, ORDINARY THINGS AND THE IMPORTANCE OF SECOND-PERSONAL SPACE

DAVID MACARTHUR

Philosophers constantly see the method of science before their eyes, and are irresistibly tempted to ask and answer questions in the way science does. This tendency is the real source of metaphysics, and leads the philosopher into complete darkness. I want to say here that it can never be our job to reduce anything to anything...
– Wittgenstein

Liberal naturalism is a novel philosophical platform from which to think about the conceptual space occupied by the philosophically neglected realm of non-scientific non-supernatural items: that is, natural things the understanding of which goes beyond the causal or causal-functional explanations made available by the sciences. Examples of such items include people (understood as rational agents), action (in the Anscombian sense of something done under a description), art (understood as having intentional content), reasons (understood as irreducible conceptually normative items), human history (understood as the history of human action), ordinary objects and much else. A liberal naturalist can readily admit the reality of such natural non-scientific things. These are among its “ontological commitments” in a deflationary sense of that expression not beholden to any metaphysical theory or Ontology. A key feature of the things within this category is that they admit of, or require, *non-scientific* forms of understanding. That they are not *fully* explicable in scientific terms is the reason I speak of them as “non-scientific” – which, of course, is not to deny that they are susceptible to various forms of scientific inquiry.

A large part of the motivation for this new kind of naturalism is in order to overcome *the placement problem* that afflicts scientific naturalism; or, rather, those versions of scientific naturalism that cleave closely to the world of physics or, slightly less drastically, the image of the world posited by the natural sciences collectively (including physics, chemistry, and biology). Given the dominance of scientific naturalism in contemporary Anglo-American philosophy the shift to liberal naturalism would have a significant bearing on how philosophy in this tradition is conceived and practiced.

On a standard version of the placement problem one *starts* with an exclusively scientific image of the world – more or less expansive versions of

which are possible – and a mismatch between this image and the manifest image of the world as it is recognized in ordinary thought, talk and experience. The standard naturalist supposes that only the scientific image (often thought of as “the physical facts”) issues in unproblematic ontological commitments. One then asks how various things in the manifest image that we are pre-theoretically committed to and which seem, *prima facie*, distinct from the scientific image are to be understood. Is our commitment to moral values, say, to be reduced to scientific items or regarded as a metaphysical ‘error’ (as Mackie put it); or are we to interpret our talk of moral values in non-representational terms, serving some other function in our lives than describing a region of reality (understood, recall, as the physical facts). This is what I am calling *the placement problem*.

Liberal naturalism overcomes the placement problem, so understood, by acknowledging, as a starting point at least, all the things that figure in the manifest image. So, for example, it draws attention to things that the practice of doing scientific research itself takes for granted: that there are people called scientists working in laboratories, whose results – published in books, journals, and on-line – are assessed in terms of the value of objectivity, and who contribute to public institutions such as universities. A liberal naturalist perspective on the activity of science itself will have to admit people, laboratories, books, journals, computers, the internet, rational values and institutions, amongst other things that do not figure in scientific explanations themselves.

But another placement problem might seem to be pressing at this point, one that threatens to undermine the impact of the liberal naturalist response to the traditional placement problem. Call this new placement problem *the relation problem*. We can put it this way: how do people, values and institutions and so on ‘fit with’ the scientific image of the world? The placement problem is one of *fitting into* but the relation problem is one of *fitting with*: it is not a matter of having to find a place for these things (say, moral values) within the scientific image of the world but of finding a place for them in relation to it.

The relation problem depends on supposing that the legitimacy of any non-scientific item (i.e. any item that does not figure in the scientific image) can only be earned by showing how it can be integrated into a synthetic account of things which includes the scientific image and the relevant non-scientific items. Thus many philosophers do not feel the liberal naturalism of John McDowell has progressed the debate very far when he argues in *Mind and World* that the intelligibility of reasons is *sui generis* with respect to scientific intelligibility. McDowell does not restrict his account of nature to the scientific image so he does not face the usual placement problems of a scientific naturalist with regard to the irreducibly normative reasons that are his primary concern. Reasons and their relations to one another are part of nature on his conception of it. McDowell explains that through a normal human upbringing (“Bildung”) we become aware of the reasons there are for judgments of various kinds and that that is all that needs to be said to assuage anxieties that might otherwise arise concerning what he calls “rampant Platonism”.

But even sympathetic critics have felt this “naturalized Platonist” position is unsatisfactory. For let us suppose we accept his liberal naturalist expansion of

nature to include non-scientific items such as reasons this still leaves them looking mysterious, so the thought goes, *unless and until* we can provide an account of how they relate to the scientific image. This is what I am calling the relation problem.

The present paper aims to argue that the relation problem is misconceived. A liberal naturalist has the resources to show that the problem is based on a false *prioritizing* of the scientific image, as if only scientific entities are real in some basic or underived way. The right thought is not that the natural world requires legitimation by being suitably related to the scientific image, as scientific naturalists would have it, but that the scientific image is an abstraction from a natural world which transcends it in so far as it also includes natural non-scientific items. This is not to say that specific questions cannot be raised about the relation of any given item to the scientific image but it does mean that the legitimacy of recognizing natural non-scientific items *in general* does not depend upon showing how they are caused, or constituted by, or otherwise dependent upon, aspects of the scientific image. Science is an *abstraction* from the ordinary world of our experience that has deeper roots in our worldview than any particular version of the scientific image.

I shall focus the discussion on ordinary artifacts which we observe and interact with every day. The choice of these items is strategic in that they are the least amenable to serious science. I shall consider the general question of the metaphysical legitimation of these objects in terms of the scientific image – a program which takes the relation problem seriously – in a recent paper of one of the more enlightened metaphysicians writing today, Lynne Baker, who is also a staunch critic of scientific naturalism. My aim will be to show that in demanding that ordinary objects be metaphysically legitimated in terms of the scientific image she makes a fatal concession to scientific naturalist orthodoxy.

THE METAPHYSICAL THREAT TO ORDINARY OBJECTS

Let us begin by considering chairs, such as my orange leather-bound Eames armchair, which will serve as a representative example of the ordinary things in the ordinary world of our experience. By ordinary things I primarily intend artifacts, items about which no one expects there to be any serious scientific findings (especially if we are thinking in terms of the discovery of causal laws). What I say could be extended to stones, trees, clouds, stars, animals and human beings in so far as there are intersubjective non-scientific criteria for the deployment of these concepts. Again, this is not to say these things are not amenable to scientific study; but it does imply that such science will not be conducted in terms of these concepts in their everyday connotation. But I shall leave these more complicated cases aside in the present context.

We can investigate chairs from the point of view of many and various sciences. We can study their physical and chemical composition, their economic value, their ergonomic indices, their social or cultural functions and so on. But none of these studies are required for us to admit the obvious fact that there are chairs. That there are chairs is attested to by the fact that we, ordinary masters of

the English language, correctly apply the concept of a chair to, well, those things that we recognize as *chairs*. That is to say the reality of chairs is manifest in our mutual recognition of certain pieces of furniture as chairs. And, importantly, nothing more is required for our acceptance of the reality of chairs than that. Barring exceptional cases – Is Daniel Bursset’s monumental artwork “Broken Chair” a chair? – we operate smoothly with the practice of using the term “chair” to pick out and make true statements about chairs, not to mention other kinds of utterance.

Contemporary metaphysicians see things otherwise. Most have something called an “ontology” and tend to give priority to an exclusively scientific version of it, scientific naturalists being the most conspicuous example of this trend. Such philosophers draw an invidious contrast between reality and appearance: that is, between an ontologically privileged category of things exhaustive of reality (e.g. the items recognized by physics) and a class of things, including ordinary objects like chairs, whose reality is thus called into question – since they do not figure, under their ordinary concepts at least, within the privileged class. For present purposes it does not much matter whether the metaphysician is an eliminativist about chairs or one who vindicates the reality of chairs by showing how they are reducible to, or irreducibly constituted by, the metaphysically privileged items.

The important point is that *all* metaphysical accounts share the basic methodological assumption that the reality of chairs requires some *metaphysical legitimation*, which may or may not be forthcoming. For example, Lynne Baker remarks,

We need a metaphysics of ordinary things to explain the epistemological point [that “we engage in practices which involve the re-identifying of objects over time”] and *to secure the rationality of practices requiring re-identification...* The basic reason to pursue a metaphysics of ordinary things is that appeal to ordinary things is needed for a coherent and comprehensive metaphysics *that secures the rationality of our practices and attitudes towards the things we encounter.* [italics added]

Let us acknowledge that unlike reductionists and eliminativists Baker laudibly accepts the irreducible reality of chairs and other common objects. But the present concern is one of *metaphysical legitimation*. According to Baker we need to accept the reality of ordinary things like chairs to give our metaphysics a coherence and comprehensiveness it would otherwise lack. At the same time, it is our metaphysics alone that can secure the rationality of our practices concerning the re-identification of ordinary things, including our thought and talk about chairs.

We can, if we wish, accept the first of these conjuncts but certainly ought to reject the second. If we are in the business of constructing a metaphysics – and I do not say we should be – then there is undoubted value in recognizing the reality of ordinary things in so far as metaphysics aspires to being a comprehensive account of everything there is. But the reality of ordinary things and the rationality of our practices of thinking and talking of them in no way

depends upon the support of metaphysical theory. Indeed, the direction of support runs entirely the other way. As Baker herself says, whatever coherence a metaphysics has depends on recognizing ordinary things that are taken for granted by our everyday thought and talk. The truism that there are chairs, which is readily attested to by all (or at least those of sound mind untainted by philosophy), in no way stands in need of legitimation or vindication from metaphysics – one dependent on the relation of ordinary things to the scientific image. As we will see, to suppose it does so depend – and so, to take the relation problem at face value – lands the reality of ordinary things in a great deal of trouble.

In any case naturalists, both scientific and liberal, ought to agree with Quine's claim that there is no first philosophy. It is simply no longer credible that we have available to us the imagined a priori standpoint of traditional metaphysics and epistemology from which to assess, from some supposed superior plane, the knowability of our claims and the genuineness of the realities that they speak of. Naturalism is incompatible with the pretensions of traditional metaphysics. So a naturalist has no reason to accept without further argument the assumption of traditional metaphysical theorizing that the reality of things in general stands in need of metaphysical vindication. Consequently, a contemporary metaphysician must demonstrate the need for metaphysical vindication and not take it for granted.

How do we recognize the reality of chairs then? As Wittgenstein says in a related context: "One answer would be: 'I have learnt English'." That may sound like a comically dismissive response. I shall now attempt to explain how it is not that.

THE COMMUNICATIVE BACKGROUND OF SCIENTIFIC PRACTICE

Let us begin with the supposed clash of the scientific and manifest images. For convenience I shall simply call this *the clash*. To think the clash exists and that it creates a pressing problem for us involves a number of questionable presuppositions. One is that the only successful sciences whose ontological commitments one is taking seriously are the *natural* sciences since otherwise, if one considers the ontological commitments of all the natural and social sciences then the gap between scientific and manifest images narrows considerably and threatens to disappear. While there may be no science of chairs, they would presumably be recognized in anthropology, sociology and economics, amongst other human sciences. Secondly, even if we do restrict our attention to the scientific image of the natural sciences, to suppose the clash is real is to suppose that the scientific image is constituted by only the *explicit* objects of scientific inquiry. This is to entirely overlook the objects (etc.) that are *implicit* or presupposed in the activity of doing science itself. And the aspect of this activity I would like to focus on is *communication with other scientists* since science is a collaborative activity on anyone's reckoning.

The scientist communicates with other scientists in what I shall call *second-personal space*. This is the familiar communicative realm of I-thou relationships that we are inculcated into when we learn a language and in which we address

each other with various linguistic utterances, demanding mutual acknowledgement of the normative import of what we say in addressing others and in being addressed by others in turn. Both the utterances and the concepts they deploy are normative items whose recognition and deployment as such occurs in the ordinary realm of mutual recognition made available in second-personal space. It is within this realm that *all* concepts must have a use, if they are to play any role in human knowledge and communication.

When McDowell distinguishes “natural-scientific intelligibility” from “the intelligibility something acquires when we situate it in the logical space of reasons” one of his important points is that scientific intelligibility prescind from conceptually normative relations which have their home in what he calls “the logical space of reasons” but which I prefer to call second-personal space. The embeddedness of scientific intelligibility in rational intelligibility is evident when we consider that a scientist is in the business of providing reasons to prefer one form of causal or causal-functional explanation of some phenomena to another (i.e. one form of scientific intelligibility over others).

Several forms of conceptual normativity are involved in communicative activity. One is that speech acts are “performances constitutive of changes in normative status among various members of a discursive community”. For instance, to assert that P is to commit to the truth of P and, often, to give one’s word to others that P, making oneself responsible for the correctness of P and entitling others to criticize one if one is not. Another is that their use is assessable in terms of standards for correctness that we deploy in second-personal space. For a concept to be correctly employed in a judgment is to be recognized as a reasonable employment of it under the circumstances by native speakers of the language. Like other concepts, *being a chair* admits of indefinitely many different occasion-sensitive understandings which we expect one another to normally be able to comprehend and track based on the most reasonable understanding of a speaker on a certain occasion of, say, making a judgment. Perhaps, under certain circumstances (e.g. for drinking a coffee on the balcony), an upside down milk crate counts as a chair. For different circumstances (e.g. people are coming over for dinner), not. As Wittgenstein reminds us, “agreement in judgment... is part of communication”.

Broadly speaking, science is an objective study of the causal structures of the world, say, of its general causal laws and local causal patterns and of the data of the coincidence of various features of the world that is the raw material for causal hypotheses. Being an objective study means the scientist is committed to *only* accepting data and results that conform to certain impersonal standards of identifiability and verification. The 3rd-p point of view of science puts certain limits on the concepts he is entitled to employ in his inquiries. As we know from the history of modern science the abstract mathematical description of nature and its development yielded enormous leaps forward in providing fruitful causal explanations of physical phenomena. It also seems clear that everyday language is run through with intentional and normative idioms that do not meet scientific standards of (re-)identifiability. Consequently, the acknowledgement of, and responsiveness to, the normative dimension of concept instantiation and the

normative pragmatics of speech acts belong to second-personal space, not the realm of scientific intelligibility.

Whether something is correctly described as a chair depends on the designer's intentions and the way we – or at least grown up people in modern industrial societies – normally use it to sit on, as well as its suitability for such use. The involvement of particular intentions, interests and goals in the identification and re-identification of artifacts like chairs makes them unfit for natural scientific inquiry whilst putting severe constraints on social scientific inquiry. Since being a chair is largely a matter of an everyday function then there is no deeper structure in the causal orders of the sciences to investigate. As Chomsky says, referring exclusively to the program of natural science,

It is hard to imagine how [“common-sense concepts”] could be fit concepts for theoretical study of things, events, and processes in the natural world.

On Chomsky's view chairs are not fit objects of naturalistic inquiry, on a restricted conception of that. But the natural scientist clearly accepts in his everyday life, and in conducting his scientific research, the reality of chairs and other common-sense objects even if they are not fit objects of his kind of inquiry. The identity conditions of everyday objects are looser and more interest relative than science would permit. Moreover, they differ from the identity conditions of the arrangements of the atoms of which they are composed since, e.g., the same chair will lose or gain atoms as it is knocked or painted or, in time, loses some of its parts e.g. rubber stops on the bottoms of its legs.

Another important theme of McDowell's distinction between natural-scientific intelligibility and rational intelligibility is that scientific intelligibility is an *abstraction* from a form of intelligibility that includes our everyday perceptual encounter with ordinary things such as chairs. In order for there to be a language at all we must share a common intersubjectively available world. In order for that to be so we must be in a position to identify and re-identify ordinary objects such as chairs by way of recognitional abilities operating on what we are presented with in perceptual experience. As Wittgenstein remarks, “Every language game is based on words 'and objects' being recognized again.”

The important conclusion we can draw from these considerations is that we should not say the natural scientist denies the reality of chairs and the ordinary criteria by which we identify them. How could he since he, too, is engaged in human communication? What we should say is that he *abstracts away from* this reality in order to do his work just as Descartes abstracted away from, without denying, beliefs that were, although not absolutely certain, more reasonable than not for the purposes of his special inquiry.

We are entitled to suppose, therefore, that scientists are inevitably committed to the ordinary observable world that finds expression in our shared language. This is a world that includes middle-sized dry goods such as chairs. Furthermore, scientists constantly engage with chairs in carrying out the business of science. The scientist sits on them to write his reports for example. And in order to do that he must normally be able to visually recognize them. To think that there are chairs is, to borrow a phrase of Hilary Putnam, one of myriad “ways of thinking

that are indispensable in everyday life”. This does not mean they could not possibly be dispensed with in our worldview. What it means is that to dispense with them would alter our lives beyond recognition. We would have to try to make sense of their never having been chairs through some science fiction scenario involving global hallucination for an indefinitely long time. One such scenario is this: we might somehow discover that oxygen, apart from being life-sustaining is also a powerful hallucinogen. But, of course, in that bizarre counterfactual case there would be a very great deal about the world we would then have been wrong about. We are inclined to say that that is just too far-fetched to take seriously.

This is not the kind of revision that metaphysicians, who would eliminate or ontologically reduce chairs, have in mind. They typically retain the concept of “chair” and do not suppose we need interfere with its everyday operation. The common strategy is to accept the truth of the sentence “There are chairs” whilst changing the accepted understanding of what it says. For example, they might say that the term “chair” *really* refers to a certain spatiotemporal worm or the sentence as a whole is reinterpreted as meaning that certain atoms are arranged chair-wise. Metaphysically motivated paraphrasings retain everyday truths involving the term “chair”, therefore, whilst being wildly at variance with our ordinary *understanding* of chairs. This shows up in qualifications that attach to the truth of their being chairs, for example, “Well, we rightly *say* there are chairs but *really* there are none over and above a certain arrangement of atoms”, (etc.). But it is important to see that in order that they correctly identify the metaphysically privileged items here they *must* employ the ordinary concept (hence the ordinary criteria) of a chair! The term “chairwise” would make no sense unless we understood what *chairs* are!

Even Baker, who works hard to overcome the clash by defending everyday truths about chairs and what she calls their “face-value reading”, runs into problems with ordinary understanding. All metaphysicians consider everyday objects groundless without metaphysical support; but this is no part of our ordinary understanding. Why, then, should we take it seriously? The source of the trouble is that metaphysicians take seriously the relation problem. They suppose, with Baker, that the only way of vindicating ordinary thought and talk about chairs is in terms of a metaphysical theory that shows how chairs are reducible to, or constituted by, the physical facts (perhaps reinterpreted in terms of some favoured metaphysical apparatus e.g. mereology, time-slices).

But once one supposes that ordinary thought and talk requires metaphysical guarantees then ordinary thought and talk is rendered groundless, even by those like Baker who mean to defend it. As Kant says in the preface to the *Critique of Pure Reason*, the history of metaphysics – and here one must surely include Kant’s own critical metaphysics – is “a battlefield of... endless controversies”. So we know from the history of philosophy that metaphysical theory is particularly ill-equipped to provide legitimation for anything given that it is, notoriously, a site of irresolvable disagreements. If we want to defend ordinary thought and talk the best strategy is to deny that it requires any (metaphysical) legitimation.

The best way to see this is to recognize the dependency of our philosophical and scientific inquiries on the realm of ordinary intelligibility. As we have seen, it is important to note that scientific study of something is an *abstraction* from what philosophers sometimes call “the life-world”, which in the terms we have already introduced is the ordinary world made available to us in second-personal space. For example, whilst he may not include his measuring equipment – the deflecting needles, LED displays, computer generated images, (etc.) – in his study it plays a key role in arriving at scientific results. Chairs are simply part of the presupposed world of scientific endeavour from which the scientific image is an intellectual abstraction. Wittgenstein writes,

In giving explanations, I already have to use language full-blown (not some sort of preparatory, provisional one)...

Yes, but then how can these observations satisfy us? – Well, your very questions were framed in this language; they had to be expressed in this language, if there was anything to ask!

Scientists and metaphysicians both have no alternative but to rely on “full-blown” ordinary language to raise their questions and communicate their results. But to use ordinary language “full-blown” means to accept the intersubjective world that it makes available to us. It is not abstracted “semantics” but the pragmatics of ordinary speech acts that makes explicit the features of the taken-for-granted ordinary life-world.

Philosophers of a metaphysical bent will sometimes cast aspersions on the reality of common-sense things such as chairs by analogy with witches or phlogiston. In medieval Europe it was common to suppose there were witches and yet modern science has shown this to be mistaken. There are none. Why couldn’t science similarly show that there are no chairs? But the cases are not at all analogous. Witches are women who are supposed to have supernatural powers (like flying unaided, causing magical transformations, consorting with supernatural beings etc.). Science has provided us the resources to reject supernatural entities and powers in favour of naturalistic explanations of natural phenomena. Chairs, of course, are not credited with supernatural powers so they are not subject to a similar debunking.

The phlogiston analogy is no better. The phlogiston theory, which posited unobservable combustion molecules that were supposed to be used up in combustion, was discovered to conflict with experimental evidence e.g. metals such as magnesium tended to gain, not lose, mass when they burned. The observable phenomenon of fire and other forms of combustion is better explained in terms of oxidation. But chairs are not, *pace* Quine, theoretical posits hypothesized to explain various observable phenomena. They are *part* of the observable phenomena, not something we have to infer from evidence. We are immediately aware of chairs through sight and touch, not to mention the sound they make when pushed along a hard floor.

THE DISPENSABILITY OF METAPHYSICS

Some argue that metaphysics itself is indispensable and so, since we are all metaphysicians whether we like it or not, the best we can do is to work out metaphysical principles that cohere as well as possible with our first-order claims, for instance, by appeal to the method of reflective equilibrium. Perhaps that is how Baker sees things. But while I accept the naturalness of metaphysical questioning and reflection I see no reason to think that the only, or the best, way of pursuing such reflections is to engage in *constructive* metaphysics. For philosophers to speak of ordinary people as having a “folk ontology”, as many are inclined to do, is an unargued prejudice, since it *assumes without reason* that ordinary folk are, just as such, *already* engaged in constructive metaphysics.

To say that ordinary practice presupposes an engagement with *ontology* is ambiguous. Taken in one way it is false; in another trivial. The relevant distinction is between doing or having an Ontology (with a capital “O”) from some (a priori) metaphysical standpoint and having ontological commitments in the course of our daily engagement with the environment. To constructively engage in the program of Ontology with a capital “O”, to aim to develop a comprehensive theory of Being a priori, is not an inevitable consequence of reflecting on the ordinary world. There is no requirement that we attempt to answer the question “What are *all* the things that exist?” for example. Indeed without a specific context of inquiry this question makes no more sense than “What are all the things that are true?”. In neither case are we given any principle to use to count *things* nor are we provided any indication of how to understand the unbounded term “all”. It matters not that metaphysicians attempt to respond to these problems. What they must show is that we are all required to take them seriously.

At one point Baker attempts to show that ordinary practice is committed to Ontological theorizing by claiming that we, in our ordinarily lives, are committed to the irrational view that “ordinary objects [such as tables] are *really just* sums of particles” (23, italics added); and that, without metaphysical support, this is an unstable position that threatens the rationality of our practices. But the woman in the street, let us say, need be credited with no such metaphysical position. And if someone is so incautious as to say that, for example, Rodin’s “The Thinker” is really just a lump of bronze they can be reminded that composition is not identity with the aid of some simple thought-experiments (e.g. would losing a flake of metal from the base really destroy the sculpture?). Ontology is an otiose project from the point of view of the rationality of ordinary and scientific forms of intelligibility.

Our everyday lives do not involve any implicit commitment to a “folk ontology” if that means, to borrow a phrase of Baker’s, “[a] complete inventory of what exists” (6). What our lives do commit us to is occasion-sensitive ontological commitments. Talk of occasion-sensitivity is appropriate for several reasons: 1) as we have seen, we need not engage with the problematic philosophical question of a *complete* inventory of things; also, 2), as Putnam has

taught us, there are different empirically equivalent ways of counting objects in certain theoretical contexts – a doctrine he called conceptual relativity; and 3) we may, or may not, decide to withhold ontological commitment in certain areas, not sure what to say, for example, if we are asked an ontological question like “Are there numbers?”

But with regard to our primary topic of everyday artifacts, we show in our thought, talk and actions that we accept such things as making up at least a part of the world. As Wittgenstein says,

My life shows that I know or am certain that there is a chair over there, or a door, and so on. – I tell a friend e.g. "Take that chair over there", "Shut the door", etc. etc.

The reality of chairs is manifest in our lives. They are indispensable to our lives as we know them. This is not to say that it is impossible that we are wrong about there being chairs. It's just that it would take an unforeseeable and massive alteration in our view of things to make it remotely credible. So much so, in fact, that it is hard to imagine a future in which it would not be more reasonable to accept the present and past reality of chairs no matter what happens.

Liberal naturalism makes available the realm of second-personal space for unprejudiced philosophical study. It is in second-personal space that we address each other and draw each other's attention to the same recognizable items in our environment thereby establishing, through mutual acknowledgement, an intersubjective world together – not that we create the world but, rather, that we identify worldly things in terms of ordinary criteria for concepts of everyday objects that native language speakers know how to deploy. Communication depends on our being able to rely on our normal capacity to be able to recognize speech acts and the speakers who produce them by way of perception. It also depends on our being able to identify and track individual objects and artifacts by appeal to ordinary criteria of identification and re-identification. The reality of ordinary things is thus a precondition of both scientific and metaphysical inquiry and does not await their discoveries.

THE POVERTY OF ANALYSIS

DAVID PAPINEAU

INTRODUCTION

Many different ideas parade under the banner of philosophical naturalism. One is a thesis about philosophical method. Philosophy investigates reality in the same way as science. Its methods are akin to scientific methods, and the knowledge it yields is akin to scientific knowledge. This ‘methodological naturalism’ is to be distinguished from ‘ontological naturalism’ understood as a general view about the contents of reality. Ontological naturalism maintains that reality involves nothing more than the entities studied in the natural sciences and contains no supernatural or transcendent realm. While both ontological and methodological naturalism claim a species of affinity between philosophy and science, the two doctrines are largely independent.

Part of the task in understanding these matters is to bring definition to this pair of naturalist doctrines. A surprisingly wide range of philosophers wish to style themselves as naturalists, and by no means all understand either the methodological or ontological commitments of naturalism in the same way. My focus in this paper will be on methodological naturalism. I shall aim to refine and defend methodological naturalism as a thesis about philosophical method. The ontological dimension of naturalism will not feature in what follows.

Methodological naturalism asserts that philosophical investigation is like scientific investigation. Clearly more needs to be said before we can subject this claim to serious assessment. Nobody can doubt that the two enterprises are similar in some respects (both aim for precision and truth, say) and different in other respects (philosophers don’t use particle accelerators). If methodological naturalism is to have any significant content, it needs to be specified in what respects philosophical and scientific methods are supposed to be alike.

I am going to argue that philosophy is like science in three interesting and non-obvious ways. First, the claims made by philosophy are synthetic not analytic: philosophical claims, just like scientific claims, are not guaranteed by the structure of the concepts they involve. Second, philosophical knowledge is a posteriori not a priori: the claims established by philosophers depend on the same kind of empirical support as scientific theories. And finally, to complete the traditional trio, the central questions of philosophy concern actuality rather

¹ Originally published in *Proceedings of the Aristotelian Society, Supplementary Volume*, Vol. 83 (2009). Reprinted by courtesy of the Editor of the Aristotelian Society © 2015.

than necessity: philosophy is primarily aimed at understanding the actual world studied by science, not some further realm of metaphysical modality.

I do not intend these claims in a revisionary spirit. I am not recommending that philosophers start doing something different. Here I diverge from other philosophers in the methodologically naturalist camp who take their position to require a shift in philosophical method—philosophers should get out of their armchairs and become more involved with active scientific research. This is not my view. When I say that philosophical investigation is akin to scientific investigation, I am not urging philosophers to change their ways. I think that most philosophy is just fine as it is, including philosophy that sticks to traditional methods of abstract theorizing, argument, and reflection on possible cases. My aim is to show that philosophy of this kind is already akin to science, not that it needs reforming in order to become so.

In what follows I shall avoid offering any positive characterization of philosophy, and in particular of what makes it different from science. For what it is worth, I do have some views about this. If pressed, I would say that philosophy is characteristically concerned with theoretical tangles. It deals with issues where deep-seated assumptions pull us in opposite directions and it is difficult to see how to resolve the tension. Because of this, the gathering of new empirical data is often (though by no means always) of no help in resolving philosophical problems. The characteristic philosophical predicament is that we have all the data we could want, but still cannot see how to resolve our theoretical problems.

Still, as I said, I am not going to commit myself to any positive characterization of philosophy. My argument does not need one. My intended subject matter is philosophy as it actually is, not a hypothetical philosophy that fits some set of prior specifications. Of course, this sociological dimension means that my claims are strictly speaking hostage to the activities of any philosophical eccentrics or extremists who deviate from my account of philosophical practice. But I hope that readers will understand my claims sympathetically in this respect. I don't want to show that everybody who has ever called themselves a 'philosopher' vindicates my claims about the nature of philosophy. It will be quite enough if I can establish my theses for those kinds of philosophy that most of you regard as mainstream.

Before proceeding, I need to qualify my claims in another respect. They do not apply equally straightforwardly to all philosophical subject matters. The areas that fit my claims best are the 'theoretical' branches of philosophy, including metaphysics, philosophy of mind, philosophy of language, and epistemology. Things become more complicated when we are dealing with areas of philosophy that trade in normative claims, or mathematical claims, or logical or modal claims. Part of the difficulty here is that the contents of these claims are themselves matters of philosophical debate, and so any attempt to show that they fit my theses about the nature of philosophy will itself become embroiled in these debates. As it happens, I think that most of the spirit of my theses about the nature of philosophy applies to these claims too, give or take a bit. But to show this would require far more space than I have available here. For present

purposes it will be enough if I can show that my theses apply to the more easily interpretable claims of theoretical philosophy.

In what follows, I shall devote most of my attention to my first thesis. The next four sections will be about the synthetic nature of philosophical claims. After that, I shall devote my final two sections to the issues of a posterioricity and modality.

THEORIES AND CONCEPTS I

It might seem that my account of philosophy falls at the first hurdle, at least in so far as it is intended as non-revisionary. What about the many philosophers who proclaim themselves to be concerned with the analysis or explication of concepts? A wide and varied range of contemporary philosophers describe their own philosophical practice as in large part concerned with the elaboration of conceptual truths. Does this not immediately belie my first thesis that philosophy as it is currently practised deals with synthetic rather than analytic claims?

I say that that these philosophers misdescribe their own practice. They may claim that they are concerned with conceptual truths, but they are wrong. When we look more carefully at what they actually do, we can see that they are in fact concerned with synthetic and not analytic matters. Indeed their claims about their practice are not even supported by everything they say they do. I shall show that when these philosophers go on to fill out their account of philosophy, their own characterization of their practice is perfectly consistent with my first thesis.

Anybody who thinks that there are conceptual truths to be uncovered must suppose that the relevant concepts have some kind of structure. They must be constitutively linked to other concepts in such a way as to place constraints on their proper application. The idea is then that this structure can be uncovered by reflection and analysis, perhaps including reflection on what we would say about a range of possible cases.

An initial question to ask about this kind of putative conceptual structure is how it relates to theories involving the relevant concepts. By ‘theories’ I mean sets of claims with synthetic consequences. A simple theory of pain in this sense would be constituted by the two claims that (a) bodily damage typically causes pains and (b) pains typically cause attempts to avoid further damage. For note that together these two claims have the manifestly synthetic consequence that bodily damage typically causes attempts to avoid further damage. We can take it that everyday thought endorses theories like this about a wide range of philosophically interesting topics, including not only mental kinds like pain, but also such categories as persons, free will, knowledge, names, and so on—after all, this is simply to assume that everyday thought includes various synthetic assumptions about these kinds.

It is widely supposed that there is a close connection between everyday concepts and everyday theories. But there are different views about the nature of this connection. In this section and the next I shall distinguish ‘verificationist’ from ‘descriptivist’ accounts of the connection between concepts and theories.

As we shall see, neither account lends any support to the thesis that philosophy is centrally concerned with analytic truths.

Let me start with the verificationist account. This assumes that possessing a concept is a matter of being disposed to use that concept in a certain way. In particular, it is a matter of applying the concept in response to perceptual experiences and other judgements, and of drawing further inferences in turn from judgements involving the concept.

Given this account of concepts, which concepts a thinker possesses will depend on what theories that thinker accepts. This is because accepting a theory affects your dispositions to apply the concepts it involves. For example, if you accept the phlogiston theory of chemistry, then you will hold that burning causes air to become saturated with phlogiston, that dephlogisticated air is easily breathable, and so on. Similarly, if you accept the baby theory of pain offered above, then you will be disposed to hold that those with bodily damage are in pain, and that those who are in pain will engage in avoidance behaviour. From the verificationist perspective, then, your commitment to these theories determines your concepts phlogiston and pain. Since the theories affect your dispositions to apply the concepts, they determine your concepts themselves.

Now, one issue which arises at this point is how much of accepted theory is supposed to make such a constitutive contribution to concepts? Do all accepted assumptions make a difference, or only some distinguished subset—and if the latter, what distinguishes this subset? However, we can by-pass these familiar questions here. The points I now want to make are quite orthogonal to this issue. They will apply to any view that takes the acceptance of sets of synthetic claims to affect concepts, however those claims might be identified.

A more basic issue is whether it makes sense to suppose that the mere possession of a concept can require a thinker to embrace synthetic commitments. Some of you may suspect that there must be something amiss with an account of concepts which implies this. However, not all philosophers share this worry. Robert Brandom, for instance, does not. He is insistent that concept possession incurs synthetic commitments. For example, after discussing Michael Dummett's example of the concept Boche, Brandon says that this

'... shows how concepts can be criticized on the basis of substantive beliefs. If one does not believe that the inference from German nationality to cruelty is a good one, then one must eschew the concept Boche' (Brandom 1994, p 126).

Again, a page later, he explains

'The concept temperature was introduced with certain criteria or circumstances of appropriate application and with certain consequences of application.... The proper question to ask in evaluating the introduction and evolution of a concept is... whether the inference embodied... is one that ought to be endorsed' (Brandom 1994, p 127).

This account of concepts plays an important part in Brandom's understanding of the philosophical enterprise. Brandon takes philosophy to be centrally concerned

with the explication of concepts. But for Brandom this is not a merely descriptive enterprise. Since concepts carry synthetic commitments, it is possible to criticize concepts on the grounds that these commitments are unwarranted. Brandom is quite explicit about this:

‘I see the point of explicating concepts rather to be opening them up to rational criticism... Defective concepts distort our thought and constrain us by limiting the propositions and plans we can entertain... Philosophy, in developing and applying tools for the rational criticism

of concepts, seeks to free us from these fetters, by bringing the distorting influences out into the light of conscious day, exposing the commitments implicit in our concepts as vulnerable to rational challenge and debate’ (Brandom 2001, p 77).

The notion that concepts have synthetic implications and are therefore open to criticism is not peculiar to Brandom. It is a commonplace of much discussion of the role of concepts in philosophy. Thus in a recent discussion of philosophical intuitions Alvin Goldman asserts that

‘A concept that embeds a bad theory is of dubious worth’ (Goldman 2007, p 22).

Again, to take just one further example, in a recent paper on moral concepts we find Richard Joyce arguing that

‘Sometimes discoveries lead us to decide that a concept (e.g., phlogiston or witch) is hopeless; sometimes we prefer to revise the concept, extirpate the problematic element, and carry on much as before’ (Joyce 2006, p 142).

I alluded a moment ago to the oddity of a view of concepts on which the mere possession of a concept can incur synthetic commitments. In fact there are further aspects of the verificationist approach that should make us even more suspicious of its account of concepts. For a start, verificationism implies that theoretical change inevitably leads to conceptual change. If you alter your theoretical assumptions involving some concept, perhaps because empirical evidence has shown that these assumptions are mistaken, then you will change your dispositions to apply that concept—and so, according to verificationism, will end up with a new concept. ‘Meaning incommensurability’ then quickly follows: adherents of different theories must mean different things even when they use the same words, and so cannot communicate with each other in a common language. In the extreme case, this implies that those who reject the ontological commitments of some theory cannot use the language of that theory to convey this. Since I do not accept the phlogiston theory, I cannot mean the same by ‘phlogiston’ as the theory’s adherents, and so cannot communicate my disagreement to them by saying ‘There is no phlogiston’.

For my money, these points are enough to discredit the verificationist account of the relation between concepts and theories. Still, I do not need to take a stand on the nature of concepts here. This is because I have no objection to what verificationists like Brandom say about philosophical practice itself, as opposed to their funny way of thinking about concepts. Brandom says that

philosophy is concerned with concepts, and then explains that for him this means that philosophy should identify the synthetic assumptions that guide our use of concepts, and criticize these assumptions when necessary. This vision of philosophical practice is entirely in accord with my first thesis that philosophy is concerned with synthetic claims.

When philosophers like Brandom say that they are explicating concepts, an unwary audience might conclude that this means that that they are not concerned with synthetic matters. But by this conclusion is belied, not only by their philosophical practice, but also by their official explanation of this practice. If the possession of concepts requires commitment to synthetic claims, and explication of these concepts involves the assessment of these claims, then there is no difference between conceptual explication and ordinary synthetic theorizing.

THEORIES AND CONCEPTS II

Even if we reject verificationist thinking, there may still be a close connection between concepts and theories. Suppose that we dismiss the notion that concept possession hinges on dispositions to apply concepts. Then our concepts will not depend on which theories we accept. But they may still depend on which theories we understand.

To see how this might work, suppose that $T(F)$ is some synthetic theory involving the concept F . Then it is open to us to regard the concept F as having its reference fixed via the description ‘the Φ such that $T(\Phi)$ ’. That is, F can be understood as referring to the unique Φ that satisfies the assumptions in T , if there is such a thing, and to fail of reference otherwise. In this spirit, we might regard pain as referring to the mental state, if there is one such, that is typically caused by damage and gives rise to avoidance behaviour, and phlogiston as referring to the substance, if there is one such, that is emitted in combustion and absorbed during chemical reduction; and so on.

On this descriptivist account, there is still a close connection between concepts and theories. But your concepts no longer depend on which theories you accept. Which theories you accept will of course affect your dispositions to apply concepts. But for non-verificationists this won’t make a difference to the concepts themselves. Even though I reject the phlogiston theory, and so apply the concept phlogiston quite differently from the eighteenth-century chemists who endorsed the theory, this doesn’t stop me having the same concept as they had. For we can all understand the concept phlogiston as equivalent to the relevant description—the putative substance that is emitted during combustion and absorbed during reduction—independently of our divergent views as to whether this description is satisfied.

In line with this, note that on the descriptivist account of concepts no synthetic commitments are incurred by the mere possession of a concept. Somebody who possesses a concept F defined by some theory T will be committed to the ‘Carnap sentence’ of the theory—if $(E\Phi)(T(\Phi))$, then $T(F)$ —but this claim will be analytic not synthetic. For example, if you have the concept phlogiston you will be committed to the relevant analytic claim, that if

there is a substance emitted during combustion and absorbed during reduction, then it is phlogiston. But you needn't thereby be committed to the synthetic commitments of the phlogiston theory itself.

From the perspective of this approach to concepts, the original theory $T(F)$ can be decomposed into the analytic Carnap sentence and the synthetic 'Ramsey sentence' of the theory— $(E\Phi)(T(\Phi))$. The Ramsey sentence expressed the substantial commitments of the theory—there is an entity which...—while the Carnap sentence expresses the definitional commitment to dubbing that entity F . The original theory framed using the concept F is thus equivalent to the conjunction of the Ramsey and Carnap sentences.

This understanding of the relation between theories and concepts informs an influential contemporary vision of philosophical practice, inspired originally by the work of David Lewis and more recently codified by Frank Jackson (1998). As conceived by Jackson, philosophy proceeds in two stages. The first stage involves the identification and articulation of folk concepts. Here the aim is to figure out how everyday thought conceives of free will, mental states, persons, moral value, and other important philosophical categories. At this stage we will use traditional methods of conceptual analysis and reflection on possible cases. Then, once we have analysed such everyday concepts, we can turn to our most serious theories of the world to investigate what satisfies them. This second stage will involve synthetic claims about the underlying nature of reality—we will look to physics and any other basic sciences to inform us about possible candidates which might realize our everyday concepts. But while this second stage appeals to synthetic knowledge, it depends essentially on the first analytic stage, where the identification of everyday concepts plays an essential role in setting the agenda for further philosophical investigation.

Thus Jackson:

'What then are the interesting philosophical questions that we are seeking to address when we debate the existence of free action and its compatibility with determinism, or about eliminativism concerning intentional psychology? What we are seeking to address is whether free action according to our ordinary conception, or something suitably close to our ordinary conception, exists and is compatible with determinism, and whether intentional states according to our ordinary conception, or something suitably close to it, will survive what cognitive science reveals about the operations of our brains' (Jackson 1998, p 31, his italics).

One worry about this programme is whether the relevant concepts really have the requisite descriptive structure. Strong externalists about content will doubt that there are any analytic assumptions involving free will say, or person, that you must be committed to if you have these concepts, let alone assumptions that will uniquely identify the referents of these concepts. (Cf Williamson 2007, ch 4.)

Another worry, which arises even if we reject strong externalism, relates to the familiar question of which everyday assumptions play a definitional role. As before, are all assumptions to be included, or only some distinguished subset—and if the latter, what marks the distinction?

I think that these are serious worries, but I shall not press them here. This is because I think I can show that, even if there are analytic truths of just the kind that that Jackson supposes, they are of no significance to philosophy.

Jackson says that everyday concepts set the agenda for further metaphysical investigation. It is because everyday thought conceives of free action, and intentional states, and so on, in such-and-such ways that we philosophers are prompted to probe the nature of those things that fit these specifications.

But why think of the matter in this way? Doesn't it make far more sense to suppose that it is the synthetic theories implicit in everyday thought that raise the initial philosophical questions, not the mere analytic commitment to concepts. Even after we allow that everyday thought is indeed structured as Jackson supposes, the natural assumption is surely that it is the synthetic Ramsey sentences that matter to philosophy, not the analytic Carnap sentences. What makes philosophers interested in investigating further is the pre-theoretical supposition that there are entities fitting such-and-such specifications, not just the hypothetical specification that if there were such entities, then they would count as free actions, or intentional states, or whatever.

The point is most easily brought out by considering cases where current everyday thought endorses the definitional Carnap sentence involving some concept, but not the substantial Ramsey sentence. I think, and so do all of you, that if there is a category of women who ride on broomsticks, cast spells, and enter into pacts with the devil, then these women are witches. But of course none of us think that there is a real kind of this sort, and so have no inclination at all to conduct metaphysical investigations into its nature. Again, to take a somewhat more serious example, we can all agree, I take it, that if there are entities that are conscious, separable from bodies, and can survive death, then those things are souls. But only those few among us who think that there actually are souls will have any motive to probe their metaphysical nature further.

The point is that concepts themselves are ontologically non-committal. The mere possession of concepts carries no implications at all about the contents of reality, and so cannot point the way to further investigations, in the way that substantial synthetic claims can.

I am very much in favour of the idea that much philosophy involves subjecting everyday ideas to serious scrutiny. All of us, philosophers included, acquire much of our understanding of the world from the everyday culture in which we grow up. Some of this everyday lore is sound, and some is not. If we are serious about our understanding of the world, we need to examine the assumptions that we acquire from everyday thinking, and see how many of them stand up to serious examination. But none of this is anything to do with concepts. Since concepts on their own are non-committal about reality, they cannot lead us astray. But the synthetic commitments of everyday thought can, and so do need to be properly examined.

When Jackson and others who subscribe to his programme actually address serious metaphysical issues, they of course proceed in just the way I am advocating. That is, they take cases where everyday thinking commits us to substantial assumptions about the contents of reality, and ask whether these

assumptions are sustainable. To this extent, I would say that their official account of what they are doing is belied by their actual practice. Officially they say they start with concepts, but in fact they start with theories.

Moreover, even the official account of what they are doing is not always strictly maintained. The difference between concepts and theories is not always respected. So in a number of passages. Jackson talks about the initial exploration of folk ideas as a matter of identifying theories rather than concepts.

For example:

‘... my intuitions reveal my theory of free action.. , your intuitions reveal your theory... to the extent that our intuitions coincide with those of the folk, they reveal the folk theory.’ (Jackson 1998, p 32)

And later we find him saying that

‘My intuitions about which possible cases to describe as cases of K-hood... reveal my theory of K-hood’ (op cit, p 37).

As I have said, I am all in favour of beginning philosophical investigation with everyday theories. But this is not the same as beginning with mere concepts. Theories involve significantly more than concepts, as is shown by the cases of witches and souls, where we have the concepts but not the corresponding theories.

THE METHOD OF POSSIBLE CASES

My thesis that philosophy deals in synthetic claims might seem to be inconsistent with one salient feature of philosophical practice. Philosophers characteristically test philosophical claims by considering whether counterexamples are in some sense imaginable. At first pass, this certainly seems to support the view that philosophical claims are conceptual in nature. Imagination can plausibly show us whether or not certain situations are conceptually possible, but presumably not whether they are actual. Correspondingly, it looks as if imagination can usefully test claims about what is conceptually required, but not about what actually occurs.

For example consider Gettier’s demonstration that knowledge is not true justified belief. Gettier showed us how to construct possible cases in which people have true justified beliefs, but are not knowers (because, roughly speaking, the truth of their belief is accidental relative to their method of justification). Surely this shows that the philosophical claim being tested is that true justified belief conceptually requires knowledge. Otherwise how could the mere conceivability of counter-examples disprove it?

Again, consider Kripke’s demolition of the descriptive theory of ordinary proper names. Kripke invited us to consider possible cases in which someone (Schmidt, say) satisfies all the descriptions associated with some name (‘Gödel’) yet is not the bearer of that name (because he is not the causal origin of its use). Here too it looks as if the mere conceivability of a counterexample is enough to discredit the thesis of interest, and thus that this thesis must be conceptual in nature.

One possible naturalist response would be to reject the method of reasoning by possible cases. Since philosophy is concerned with synthetic claims, just like the sciences, it can't possibly make progress just by reflecting on what is conceptually possible. Instead philosophers should get out of their armchairs and engage directly with experimental and observational findings.

This is not my view. I take it to be uncontentious that Gettier's and Kripke's thought-experiments led to genuine advances in philosophical knowledge. More generally, I regard reflection on possible cases as a highly fruitful mode of philosophical investigation. As I said at the beginning, I am not proposing any revisionary account of philosophical practice. From my point of view, the methods that philosophers use are just fine, including the method of reflection on possible cases. So instead of rejecting armchair reflection, I am going to argue that armchair methods provide more than purely conceptual information and so can play a part in the assessment of synthetic claims.

The obvious comparison here is with thought-experiments in sciences. Many important advances in science have been prompted by pure reflection on possible cases. Famous examples include Archimedes on buoyancy, Galileo on falling bodies and the relativity of motion, Newton's bucket experiment, Maxwell's demon, and Einstein on quantum non-locality. Cases like these certainly suggest that armchair reflection can be relevant to establishing synthetic claims.

Scientific thought-experiments display a range of different structures. Let me focus on one of the simpler cases—Galileo's analysis of falling bodies. According to the Aristotelian orthodoxy of Galileo's time, heavier bodies fall faster than lighter ones. Galileo asks his readers to consider what will happen if a lighter body is tied to a heavier one by a piece of string (Galileo 1638). Since the Aristotelian theory says the lighter body will be inclined to fall more slowly than the heavier, it follows that the lighter should slow the heavier down when joined to it. But by the same coin the compound body consisting of the two tied together is heavier than the two individual bodies, and so should fall faster than both. The Aristotelian theory is thus shown to be inconsistent. Moreover, it looks as if the only consistent account will have the compound body falling at the same speed as the individual components, which implies that speed of fall is independent of weight.

In this kind of case it is clear that the relationship between weight and speed of fall is a synthetic matter. Concepts cannot guarantee anything this relationship. How then can armchair reflection show us what to think? The answer must be that armchair reflection is showing us more than that certain scenarios are conceptually possible. Of course, it can't show that there are any actual cases in which a compound body falls at the same speed as its components. Galileo didn't create a real case of two bodies tied together just by thinking about it. Still, Galileo didn't need an actual case to disprove the Aristotelian theory. If we construe that theory as saying that the faster fall of heavier bodies is required by the laws of nature, it will be enough for Galileo to show that a case of a heavier body falling at the same speed as a lighter one is consistent with the laws of nature. And that is just what Galileo does. He asks us to consider a manifestly naturally possible scenario in which two bodies are tied

together, and then judges that in such a case the laws of nature will lead the compound body to fall at the same speed as its components.

Obviously, the crucial step here is played by Galileo's intuition that a compound body will fall at the same speed as its components. And this is clearly a synthetic intuition, by no means guaranteed by the concepts it involves. That is why it can overturn the synthetic Aristotelian theory.

I want to suggest that philosophical thought experiments have the same structure. Explicit philosophical theories about the requirements for a thinker to know something, or for a thing to bear a name, (or for someone to have acted freely, or for one person to be the same as another,...) are synthetic claims about the relevant categories. Philosophers then test such synthetic proposals against their intuitions about possible scenarios. Thus Gettier appealed to the intuition that a belief whose truth is accidental relative to its method of justification is not knowledge; Kripke appealed to the intuition that something that is not the causal origin of a name is not its bearer; and so on. On my account, all these intuitions are synthetic claims about the relevant kind of scenario. This is why they have the power to discredit the initial philosophical theories.

From this perspective, there is nothing in the method of reasoning about possible cases to undermine the idea that philosophy is concerned with synthetic claims. It is simply a technique that enables us to counter the synthetic theories proposed by philosophers by the synthetic intuitions elicited by thought experiments.

There is one respect in which this account of thought experiments may be an oversimplification. I have suggested that thought experimental intuitions manifest certain general principles, such as that an accidentally true believer isn't a knower, or that the causal origin of a name is its bearer, and so on. However Tim Williamson has pointed out that such general claims are arguably more than the thought experiments committed us to (2007, Ch 6). For example, in order to disprove the tripartite analysis of knowledge, Gettier only needed the particular counterfactual claim that, in the most obvious understanding of his scenario, the relevant thinker would not be a knower. There is no need to suppose that any thinker satisfying the explicit specifications of his scenario would fail to know, still less to suppose some still more general principle as that 'all accidentally true believer aren't knowers'. For Williamson, philosophical thought experiments thus appeal only to our ability to reason counterfactually, and do not demand any grasp of general principles.

I am happy to agree that counterfactual reasoning is enough for thought-experimental purposes, and correspondingly that is in by no means mandatory to suppose that general principles lie behind the relevant intuitions. Even so, I would like to continue working on the assumption that thought experiments display general principles. This may be an oversimplification, but I don't think it is too far from the truth. We may not fully understand counterfactual reasoning, but it is clear that it is strongly constrained by general claims about the working of the world. Williamson alludes to the role of imagination in counterfactual reasoning (2007, chs 5-6). But when I think about what would happen if I had dropped a vase, say, I do not imagine every outcome that is permitted by the concepts involved, such as that the vase floats gently onto the table. Rather I

consider just those outcomes that are consistent with some such synthetic general claim as that heavy bodies fall rapidly when unsupported. Perhaps this general claim as just formulated is more precise than anything that governs our counterfactual thinking. Still, it seems clear that our counterfactual thinking must be informed by some such principle. In line with this, I shall continue to assume that the intuitions in philosophical thought experiments are informed by general principles. Attempts to state these principles explicitly may inevitably lead to oversimplification, but I propose to overlook this in the interests of facilitating investigation into their nature. (In what follows I shall use ‘intuition’ to refer to both the general principles informing our counterfactual reasoning and the specific judgements about counterfactual situations that issue from them. When the distinction matters I shall draw it explicitly.)

THE ENCAPSULATION OF ASSUMPTIONS

There is an obvious objection to my proposed analogy between philosophical and scientific thought experiments. Consider Galileo’s thought experiment again. The crucial intuition was that tying two bodies together won’t make any difference to their speed of fall. Now, it is clear that this conjecture is hostage to further empirical investigation. It may strike us as obvious that Galileo is right, but even so empirical observation remains the ultimate test of his intuition. Galileo is in effect hazarding a guess—albeit a highly informed guess—as to the synthetic facts, and the final arbiter of this guess must be real observations. We need to find some actual bodies that are tied together see how they fall. Either they will conform to Galileo’s intuition, or they won’t. And both options are clearly left open by the terms in which the issue is posed.

Things seem rather different in philosophy. In the Gettier thought experiment, for example, the analogous intuition was that a belief isn’t knowledge if its truth is an accident relative to its method of justification. But there seems no analogous room to check this intuition against real cases, by seeing whether or not actual thinkers with such accidentally true beliefs are knowers. For we already know what we will say about any such cases—namely, that these thinkers are certainly not knowers. The reflection involved in the philosophical thought experiment is itself enough to tell us what we will judge in any similar real situation, and thus to rule out any possibility of observing someone who is an accidentally true believer yet a knower. The same seems true of philosophical thought experiments in general. Take the Kripke case. We don’t need to find any real cases of names whose original bearers don’t fit the associated descriptions, in order to check whether or not the names really do name the original bearers. For again, we already know what we will say about any real such cases—namely, that the names apply to the original bearers even if they don’t satisfy the descriptions. And this again rules out any possibility of observing a name which turns out to refer to the satisfier of associated descriptions rather than the original bearer.

In short, the intuitions in play in the philosophical thought experiments don’t seem to be falsifiable in the way they ought to be if they were synthetic claims. On the contrary, their inviolability to any observational falsification seems to argue strongly that they are analytic. And this would then imply that the

philosophical thought experiments are serving to manifest the structure of our concepts, rather than to draw out our implicit empirical opinions.

However, this is not the only way of seeing the matter. An alternative is to hold that the relevant philosophical intuitions are synthetic, but encapsulated in the cognitive systems that make judgements about such categories as knowledge, names, free will, persons and so on. By way of analogy, consider the way that the human visual system detects the edges of physical objects by registering sharp changes in intensity in the visual field. We can think of the visual system as embodying the implicit ‘assumption’ that intensity changes are due to the edges of physical objects. This assumption is then ‘encapsulated’ in the sense that the visual system will continue to embrace it even in cases where we are personally aware that the intensity changes are due to something else, as when we are viewing the surface of a photograph.

Because of this, it is inevitable that we see intensity changes as edges. And this means that our visual system is never going to deliver intuitive particular ‘judgements’ that falsify the intuitive general ‘assumption’ that all intensity changes are due to object edges. There is no possibility of a visual observation of sharp intensity changes that are not seen as edges. Still, it is clear enough that the assumption that all intensity changes are due to object edges is akin to a synthetic rather than an analytic claim. Its approximate truth is not due to the structure of its content, but to the fact that most intensity changes in the actual world are due to the edges of physical objects.

I would like to say the same about the general intuitions which guide us in making particular judgements about knowledge, names, persons, free will and so on. The sub-personal cognitive mechanisms responsible for such judgements are not well understood, as is evidenced by the difficulty philosophers have in identifying the principles on which they operate. But it is clear enough how they must work: they take in information which do not presuppose the relevant categories, and use it to arrive at judgements about who knows what, and which words name which things, and when someone is the same person as someone else, and so on. I want to suggest that the particular intuitions displayed in philosophical thought experiments manifest the implicit general ‘assumptions’ on which such mechanisms hinge, in the way that the visual system hinges on the ‘assumption’ that intensity changes are due to object edges.

This is why there is no question of any direct judgements about particular cases falsifying such ‘assumptions’. If my judgmental procedures decide who is a knower by assuming *inter alia* that accidentally true believers are not knowers, then clearly there isn’t any question of my meeting up with a case where I judge such an accidentally true believer to be a knower after all. Again, if my judgmental procedures decide what things bears some name by noting the causal origin of the use of the name, then I’m not going to come across cases where I judge that some name is borne by something other than its causal origin. But this impossibility of direct falsification does not mean that the relevant general assumptions are analytic. They may yet have a substantial synthetic content, like the visual system’s assumption that intensity changes are due to object edges.

Some readers may be feeling that I have not yet established a positive case for my first thesis that philosophy deals in synthetic matters. In this section and the last I may have succeeded in showing how the importance of thought-experiments can be made consistent with that thesis. But isn't it equally consistent with the contrary thesis that philosophy is centrally concerned with analytic matters? I may have been able to concoct a story which makes philosophical thought-experiments come out like scientific ones. But isn't the more natural account still that the point of these thought-experiments is to articulate the structure of our concepts? Don't I owe some positive arguments against this natural account and in favour of the one I have contrived?

I have two responses to this line of thought. First, there are independent reasons for thinking that substantial synthetic assumptions are built into the automatic mechanisms that allow us to make particular judgements about philosophically salient categories like knowledge, names, persons, free will and so on. Judgements like these are important to us in our daily life, and it is therefore unsurprising that we should have unthinking mechanisms that allow us to form them quickly and efficiently. But it would be odd then to suppose that any inferential assumptions built into these mechanisms must be analyticities whose truth is guaranteed by the structure of their contents. The whole point of these mechanisms is to start with limited information and deliver further conclusions. It would run quite counter to this function if they were restricted to analytic inferences and precluded from engaging in ampliative ones.

My second response is that, if philosophical thought-experiments were concerned only with the elaboration of analyticities, they would be much less interesting than they are. They would tell us about the structure of our concepts, but they wouldn't help us to understand the rest of the world. (Cf Williamson 2007, 204-7.)

Recall that analytic knowledge comes in the form of conditional Carnap sentences. These simply explain that, if things satisfying certain requirements exist, then they count as such-and-suches, but analytic knowledge never deliver any categorical information about the contents of actuality. Correspondingly, the philosophical analysis of concepts may tell us that, if there is a propositional attitude that requires truth, justification, and so on, then it is knowledge—or again that, if words and things bear certain causal relations, then the words name the things.

But this seems far less than we actually get from the relevant thought-experiments. Thus I take Gettier to have shown not just that our concept of knowledge imposes a requirement of non-accidentally, but far more interestingly that this requirement is satisfied by real knowledge—that is, the state that plays an important role in the world and is displayed in many paradigm cases. Similarly, I take Kripke to have shown that not just that we conceptualize names causally, but in addition that real name-bearer pairs—all those many instances we are familiar with—are causally related.

If the assumptions manifested in philosophical thought experiments really are synthetic, then of course their contents leave it open that they may turn out to be false. They may not be directly falsifiable via a simple contrary observation, for the reasons outlined above. Still, we can imagine how more sophisticated investigations may show them to be flawed. Compare the way in which, even

though we never see sharp intensity changes as anything other than object edges, more elaborate investigation can show us that there are plenty of contrary cases. We can imagine reaching a similar conclusion about knowledge, say. We are notoriously unclear about the significance of knowledge. (Is it just that knowledge is an effective means to true belief, or because knowledge underpins a certain kind of robustness in action explanations, or because it is biologically more basic than true belief, or for some further reason?) Perhaps when we have a good answer to this question we will conclude that the principle that knowledge requires non-accidentality is a crude rule of thumb that works well enough in many cases, but on occasion leads us astray. (I doubt that it will actually turn out like this. My concern here is only to show how it is epistemologically possible.)

PHILOSOPHY IS A POSTERIORI

Let me now turn to my second thesis that philosophical knowledge is a posteriori not a priori. It might seem that this will now follow quickly, given my first thesis that philosophical claims are synthetic. How can a synthetic claim possibly be known to be true independently of experience, given that its content alone leaves it open that it might be false? But of course this is too quick. Traditional theists and transcendental idealists both take themselves to have good answers to this question. And even if we reject these particular answers, there is room for other non-experiential accounts of synthetic knowledge.

On this topic, Timothy Williamson has argued that philosophical intuitions, though synthetic, should not be counted as a posteriori (2006, 165-9, 189-90). His reason is that experience does not play a normal evidential role in generating them. We can't point to past observations of supporting instances to support such claims as that accidentally true believers are not knowers, or that names refer to their causal origins. Our route to these claims is thus clearly unlike the normal justification of synthetic generalizations by inductive or abductive evidence. (Williamson does not conclude that such philosophical judgements are a priori—he thinks the traditional contrast is not useful here. But we need not pursue this point, given that I am about to argue contra Williamson that philosophical intuitions should definitely be counted as a posteriori.)

I am in broad agreement with Williamson's view of the provenance of philosophical intuitions. They are not products of normal inductions or abductions. This accords well with my suggestion that the underlying assumptions are 'encapsulated' in the cognitive mechanisms we use to decide on cases of knowledge, naming, and so on. Of course, there is plenty of room for debate about the means by which assumptions become encapsulated in this way. Strong nativists argue that all the relevant information is encoded in genes bequeathed to us by natural selection. Others hold that the relevant assumptions are laid down early in individual development, via the influence of surrounding culture and possibly also of acquaintance with particular paradigm cases. Still, whatever the precise truth on this matter, Williamson seems right to observe that the source of philosophical intuitions is not normal inductive or abductive evidence.

Still, the source of philosophical intuitions is one thing, their justification another. Even if philosophical intuitions do not derive from experience, it may still be that they can only be justified a posteriori. This is my view. Note that in general the epistemological status of encapsulated assumptions is not high. The function of cognitive mechanisms that embody encapsulated assumptions is to deliver judgements about particular cases quickly and efficiently. Because of this, the relevant assumptions are standardly rules of thumb that work well enough in most cases but are not strictly accurate, in the way illustrated by the familiar perceptual examples. If the cognitive mechanisms behind philosophical intuitions are at all similar, we should expect encapsulated philosophical assumptions to have a similar status. They may work well enough for practical purposes, but they may not be strictly accurate and may lead us astray in certain cases. If we are to be confident about these assumptions, we will need to make them explicit and subject them to proper a posteriori evaluation.

I have already argued that it is at least epistemologically possible that there may be inaccuracies in the assumption that knowledge must be non-accidental. Nor is it hard to think of real cases of mistaken assumptions which were once encapsulated and therefore seemed immune to imaginable counter-examples. Descartes thought it unimaginable that a purely mechanical being could reason. Kant thought it unimaginable that parallel lines could meet. Many people still find it unimaginable that temporal succession could be relative, or that time could have a beginning.

The recent findings of ‘experimental philosophy’ are relevant here. They indicate that many central philosophical intuitions, including those invoked by Gettier and Kripke, are by no means universal, but rather peculiar to certain cultures and social classes. (Knobe and Nichols eds 2008.) At one level, it is not always clear what to make of these findings. Presented as a challenge to ‘conceptual analysis’, they invite the response that the variability of intuitions only establishes the philosophically insignificant point that different groups of people express different concepts by words like ‘knowledge’ and ‘name’. However, the variability of intuitions is clearly more significant if philosophical intuitions are substantial claims whose truth is not analytically guaranteed. In that case, the variability of the intuitions is in tension with their reliability. If different people have opposed philosophical intuitions, then it cannot be that intuitions of this kind are always true. This reinforces the point that an a priori provenance for philosophical assumptions does not amount to an a priori justification. As before, the justification of such assumptions requires that we subject them to proper a posteriori examination.

It may seem as if I am here backtracking on my earlier enthusiasm for armchair philosophy. Among philosophers who agree with me that philosophical intuitions are synthetic, we can distinguish two broad positions. There are those who think that philosophical intuitions are little more than manifestations of naïve folklore, and should therefore carry little weight in serious philosophical discussion. According to this point of view, philosophers should turn away from intuitions and instead engage with serious empirical theories. (Cf. Kornblith 2002, Knobe and Nichols eds 2008.) On the other side are philosophers like Timothy Williamson, and perhaps Alvin Goldman, who think that philosophical intuitions are by and large reliable, and that

the findings of the experimental philosophers are not as worrying as they appear. (For example, Williamson suggests that the special training of philosophers may make them sensitive to niceties that escape the untrained—2007, 191.) My line of argument in this section so far may seem to place me on the former side and thus against armchair investigation.

However, I think this conclusion is based on a false dichotomy. Just because I am doubtful about the authority of philosophical intuitions, it doesn't mean that I have to reject the method of reasoning about merely possible cases. Armchair thinking can be useful, even if the intuitions involved are unreliable. Go back to the idea, briefly aired earlier, that philosophy is characteristically concerned with theoretical tangles. We find our thinking pulled in opposing directions and cannot see how to resolve the tension. Often part of our predicament is that we don't know what assumptions are directing our thinking. We end up with conflicting judgements, but are unclear about what led us there. In such cases thought experiments can bring the implicit principles behind our conflicting judgements to the surface. They make it clear what intuitive general assumptions are governing our thinking and so allow us to subject these assumptions to explicit examination. Nothing in this requires that thought-experimental thinking is generally reliable. When some explicit prior theory conflicts with an intuitive judgement elicited by a thought experiment, this needn't always result in the rejection of the theory. We can also end up rejecting the implicit assumptions behind the thought-experimental intuition.

Just this pattern is displayed by some of the most famous and important thought experiments in science. Consider the 'tower argument' against Copernicanism, which appeals to the intuition that an object dropped from a moving source will be 'left behind' as it falls. Or take the Einsteinian argument against the completeness of the Copenhagen interpretation of quantum mechanics, which appeals to the intuition that spacelike separated events cannot be coordinated without a common cause. In cases like these, the assumptions generating the thought experiments eventually came to be recognized as mistaken. But this certainly did not mean that the thought experiments were worthless. Both the tower and Einstein arguments were hugely important in the history of science. But showing us which of our implicit assumptions conflicted with new theoretical ideas, they led to crucial new advances. Galileo responded to the tower argument with his innovative formulation of a principle of inertia, and J.S. Bell to the Einstein argument with his derivation of the eponymous inequality whose experimental confirmation ruled out local hidden variable theories.

It is not hard to think of similar philosophical cases. The worth of philosophical thought experiments does not always require that the intuitions they elicit are sound. In some cases, of course, the intuitions will be correct. I don't think we should really harbour any serious doubts about the Gettier and Kripke intuitions. But in other cases thought experiments can clarify the issues even if the accompanying intuitions point us in the wrong direction.

Consider the classic Lockean set-up where someone's memories are transferred to a new body. We all have an intuition that the person goes with the memories, not the old body, as evidenced by our reactions to the many fictions which trade on just this kind of scenario. But few philosophers of personal

identity would nowadays hold that this intuition is decisive in favour of Lockeanism. We need to follow through the implications of the Lockean views and assess the overall resulting theory against its competitors, and in this context the initial intuition is indecisive. But, for all that, it would be hard to deny that Locke's thought experiment has led to advances in our understanding of personal identity. Again, consider the intuition that conscious properties are ontologically distinct from physical ones, as displayed in our immediate reaction to zombie scenarios. Here too few would suppose that these intuitions are decisive in refuting physicalism. But at the same time even physicalists will allow that reflection on zombie cases has helped to clarify what is at issue in the mind-brain debate. (I shall return to this particular example in the next section.)

So my view is that philosophical intuitions do no qualify as knowledge until they have been subject to serious a posteriori assessment. Philosophers need to articulate their intuitions in order to understand the source of their theoretical difficulties. But since these intuitions are standardly nothing more the encapsulated rules of thumb we happen to have grown up with, we should not place any great epistemological weight on them until they have been properly evaluated against experience. In saying this, I do not mean to imply that all philosophical claims need to be assessed directly against specific empirical findings from empirical disciplines. A synthetic theory can be vindicated a posteriori even though it has no specific empirical evidence to call its own, on the grounds that it provides a more coherent and natural overall account than the alternatives.

As it happens, I do think that specific empirical findings bear directly on a surprisingly wide range of philosophical issues. These include not just topics from philosophy of science, such as the logic of natural selection or the interpretation of quantum mechanics, but also such central and traditional topics as the nature of causation and the relation between mind and brain. Still, I am happy to allow that there are other central philosophical issues, such as the nature of persisting objects or realism about properties, where the philosophical claims float free of any specific matters investigated by the empirical sciences.² In such cases, we will then have no alternative but to evaluate alternative philosophical positions by comparing their overall coherence and naturalness. Still, this too is an a posteriori procedure, akin to the method by which we compare alternative scientific theories that are underdetermined by the evidence. When we prefer Copernicus to Ptolemy, or special relativity with the Lorentz-Fitzgerald reworking of classical mechanics, it is not because of any specific empirical findings, but because they are more in accord with our general a posteriori understanding of the way the world works. I see no reason to doubt that the most abstract philosophical issues are to be decided in the same way.³

² But see Maudlin, 2007, who brings scientific considerations to bear even on these two topics.

³ Some readers might wish to query whether choices between underdetermined scientific theories should count as a posteriori. I think that they should (1993, Ch 5), but perhaps I can let the point pass here, and settle instead for the observation that empirically underdetermined philosophical theory-choices are made on the same grounds as scientific ones.

PHILOSOPHY AND NECESSITY

Can the account of philosophy offered so far accommodate the modal dimension of philosophical knowledge? It is sometimes said that the difference between philosophy and science is that philosophy seeks necessary truths where science trades in contingencies. (Thus Russell: '[a philosophical proposition] must not deal specially with things on the surface of the earth, or with the solar system, or with any other portion of space and time.... A philosophical proposition must be applicable to everything that exists or may exist' 1914, p 110.) This modal view of philosophy might seem to be in tension with my account of philosophy as synthetic and a posteriori. Don't we need a priori analysis to uncover necessary truths?

But of course this line of thought is far too quick. There no reason why necessities should not be synthetic and a posteriori. Empirical science provides plenty of familiar examples. Water is H₂O. Heat is molecular motion. Stars are made of hot gas. Halley's comet is made of rock and ice. All these claims are necessary, but clearly they are not knowable a priori on some analytic basis.

These claims are necessary because they use rigid terminology to report on facts of identity or constitution. All claims of these kinds are necessary, notwithstanding any synthetic a posteriori status they may have. It is a nice question, worthy of further discussion, why facts like these should count as necessary, while truths about spatiotemporal location, say, do not. But this is not the place to pursue this issue. For present purposes the important point is simply that the necessity of claims of these kinds is perfectly consistent with their synthetic a posteriori status.

The central questions of philosophy are almost entirely concerned with issues of identity and constitution. When we ask about knowledge, names, persons, persisting objects, free will, causation, and so on, we are seeking to understand the nature of these categories. We want to know whether knowledge is the same as true justified belief, whether naming involves descriptive content, whether persisting objects are composed of temporal parts, and so on. Any truths we might establish about such matters will inevitably be necessary rather than contingent, even if they are also a posteriori and synthetic.

The answers to the central questions of philosophy may be necessary, but that is no reason to suppose that philosophy is here concerned with necessity per se rather than actuality. Consider empirical science once more. As I have just observed, many of the claims established by science are necessary. But it would be odd to infer from this that empirical science is aiming to explore some wider modal realm rather than simply to understand the actual world. When science investigates the chemical make-up of water, or the composition of the stars, it is primarily concerned with how things are in this world ('with things on the surface of the earth, or with the solar system, or with any other portion of space and time...') That these discoveries have implications about the contents of other possible worlds, so to speak, is an inevitable side-effect of the content of these claims, but not something that we need regard science as actively seeking.

I say the same about the central areas of philosophy. Our primary philosophical concern is to find out about things in this world. We want to know

about such actual categories as knowledge, free will, persons, and so on—kinds that exist and make a difference in this world. Of course, given that answers to our questions will normally take the form of claims about identity and constitution, philosophical knowledge will also place constraints on what is necessary and possible. But there is no reason to regard such modal corollaries as our main aim. We are first seeking to understand this world, and are only derivatively concerned with modal matters. We want to know whether p , not whether necessarily p . That the former implies the latter does not make the latter our focus of interest, any more that my interest in whether you are 47 years old makes me interested in whether your age is a prime number.

Of course, some philosophers are specifically interested in modal questions as such. They are interested in whether necessary truths are necessarily necessary, or in whether modal claims commit us to an ontology of possible worlds, or in the connection between metaphysical and conceptual necessity, or indeed in why facts of identity and constitution but not spatio-temporal location should count as necessary, and so forth. There are certainly substantial philosophical issues worthy of serious discussion.⁴ But most central philosophical questions are not of this form. The study of modality is a specialist subject within philosophy, engendered by specific theoretical interests. There is no reason to suppose that an interest in modality infects all of philosophy, even if all philosophical claims have modal implications.

Having said this, it is worth recognizing that it is often heuristically useful to focus on modal implications, even in cases where our real interest is in non-modal matters. Given the immediate modal upshot of claims of identity and constitution, it is sometimes easier to articulate our thinking by starting with the modal consequences rather than their this-worldly counterparts. Take the relation between individual objects and their property instantiations. In the actual world there is a one-to-one correspondence between objects and sets of property instantiations. But is this a matter of identity, as in the ‘bundle theory’ of objects, or mere association? A good way to clarify our thinking on this issue is to consider the modal question of whether there could be a world in which this blue cup, say, acquired all the properties of that red one, and vice versa. To the extent this strikes us as possible, then we are thinking of objects as distinct from their property instantiations; while if it seems that this is not a real possibility, then we are identifying objects with their property instantiations.

I am not of course here suggesting that such modal intuitions are somehow a privileged route to the truth. Whether we are right to think of objects as bundles of properties, say, would remain a substantial further issue, even after modal reflection has made it clear that this is our intuitive view. The role of the modal reflection is merely to clarify the content of our intuitive commitments in cases where thinking about actuality alone leaves them unclear, not to decide the

⁴ This branch of philosophy obviously demands a qualification to the third of my initial theses—philosophers of modality are certainly concerned to understand modality per se, even if other philosophers are not. But it may still satisfy my other two theses by being synthetic and a posteriori. As before, however, we cannot expect decide these issues in the absence of an agreed view about the nature of modal claims.

substantial issues. From this perspective, modal thinking is a special case of the kind of thought-experimental reflection described in previous sections. It is a useful way of identifying the implicit assumptions that drive our reasoning. Once these assumptions have been identified, we are then in a position to subject them to serious a posteriori evaluation.

Let me conclude with one further example. Consider the relation between conscious mental properties and brain properties. Let us agree that pairs of these properties go hand-in-hand in the actual world. Still, is this association due to the identity of the relevant properties, or merely to a correlation between distinct properties? Well, ask yourself whether there could possibly be a being with all your brain properties but who lacks your conscious properties. If you think that such zombies are possible, then you must be of the view that conscious properties are distinct from brain properties in this world. Conversely, if you think that conscious properties are in actuality one and the same as brain properties, then you won't think that zombies are so much as possible.

Many recent writers look at this thought experiment differently. They think we can start with our concepts of conscious and brain states, proceed to the point that zombies are conceivable, somehow move from this to their possibility, and thence end up with the conclusion that conscious and brain properties are distinct in the actual world. (Chalmers 1996, Bealer 2002.) I don't think that this works at all. (Papineau 2007.) The interesting thing about zombies isn't that we can conceive them—after all, we can conceive lots of things that aren't possible—but that they strike us as possible. This shows us something rather surprising, namely that at an intuitive level we are all dualists about the mind-brain relation.

Of course, it is one thing to identify this intuition, and another to justify it. As I have argued throughout, philosophical intuitions need a posteriori backing before we can place confidence in them. In this case it seems clear that the a posteriori evidence counts against the intuition. (Papineau 2002, Appendix.) Still, this is not the place to pursue this issue, which is in any case independent of my present point—which is that in most familiar cases the purpose of modal reflection is not to find out about other possible worlds *per se*, but simply to clarify our pre-theoretical assumptions about the actual world.⁵

⁵ I have presented versions of this material in many places, including a seminar on the philosophy of philosophy in London in 2008, and I would like to thank all those who responded on these occasions. I can particularly remember helpful comments from George Bealer, David Chalmers, Keith Hossack, Fraser McBride, Tom Pink, Andrea Sangiovanni, Gabriel Segal, Jonathan Shaffer, Barry Smith, Stephen Stich, Scott Sturgeon, Celia Teixeira, Mark Textor, Lee Walters, Tim Williamson and Crispin Wright.

DEMYSTIFYING THE NATURE OF MORAL JUDGMENTS BY REDUCTIONIST NATURALISM

EVANDRO BARBOSA

INTRODUCTION

Currently, moral naturalism seems to be a powerful tool for solving the problem of justification. Among its supporters, David Copp offers his Society-Centered Moral Theory as a reductionist kind of naturalism, which connects realism and constructivism. Nevertheless, this seems to be a conflicting point in his work that needs to be clarified. According to Copp, the understanding of the moral landscape requires three basic interrelated points: identifying the nature of moral judgments, specifying the justification of those decisions and determining the role of moral society in this scheme. Hence, the philosopher discusses the necessity of a cognitivist theory of normative language (standard-based theory) as well as a theory of moral codes justification (society-centered theory). In 1995, Copp wrote *Morality, Normativity and Society*, in which he systematically defends his philosophical position as a naturalist, realist and constructivist. Later, in his book *Morality in a Natural World* (2007),¹ he revises his position and discusses why it would be better to explain his naturalist theory as a non-constructivist one, although he provides no definitive position on this matter. In spite of further changes and detailing, the leitmotif of his theory remains intact: the defense of a reductionist and realist moral naturalism reluctant whether to sustain a constructivist position or not. This chapter analyzes the extent to which Copp's moral theory is able to support a constructivist methodology on the basis of his realist moral naturalism without undermining his naturalist purpose of relating (and reducing) moral properties to natural ones. To determine this relationship, broad definitions – which I regard as *lato sensu* ones – of

¹ There is a doubt in his theory that will follow our discussion: “One important distinction that I need to explain is between the ‘constructivist’ version of the theory that I presented in my first book [*Morality, Normativity, and Society*] and the ‘nonconstructivist’ version that is at work in the present book [*Morality in a Natural World*]. I believe that the nonconstructivist version is preferable.” (Introduction) Recently, Copp wrote a text titled “Is constructivism an alternative to moral realism?” (2013), which is relevant from the point of view of systematization of the relationship between realism and constructivism in his theory, insofar as he says that there is no contradiction in approximating both, provided the settings are clearly determined.

naturalism, realism and constructivism² will be introduced with the purpose of showing their specificities in Copp's theory and of identifying the reason why naturalism does not convincingly connect realism and constructivism. Further, the core elements of his theory will be presented – specially his cognitivist perspective of normative language (standard-based theory) and his theory of moral codes justification (society-centered theory) based on his Society-Centered Moral Theory.

I – *LATO SENSU* NATURALISM AND MORAL NATURALISM

It is somehow difficult to specify what the term “naturalism” means to different philosophers, since a wide variety of definitions follows the huge range of areas of application. Nevertheless, one element is generally agreed upon: all models of naturalism reject the possibility of supernatural entities.³ In this sense, the naturalist is based on a metaphysical position according to which all the facts, including moral ones, are natural facts.⁴ Harman defines naturalism in a broad sense as follows:

Philosophical naturalism is a special instance of the wider conception of philosophy, taking the subject matter and *methods* of philosophy to be continuous with the subject matters and methods of other disciplines, especially including the natural sciences. From a naturalistic perspective, productive philosophers are those who (among other things) produce fruitful more or less speculative theoretical ideas, with no sharp distinction between such theorizing by members of philosophy departments and such theorizing by members of other departments.⁵

In the normative and metanormative spheres, naturalism is a title frequently used to indicate a point of view that combines moral realism and moral naturalism. This relationship defines the standard position of naturalism. Such position assumes the naturalist methodological strategy and the realist ontological condition as a powerful combination not only to support objective moral facts, but also to explain the relationship between moral properties and natural properties. In this regard, while the non-naturalist can appeal to entities or properties considered non-natural to explain the moral phenomenon, the naturalist does not support a domain consisted of non-natural or supernatural facts. Therefore, in face of questions related to the moral value and its justification basis, the naturalist answer goes through the possibility of supporting values and moral duties on scientific and naturalist grounds. Sturgeon defines the standard position of moral naturalism as involving two properties:

(a) that such ethical properties as the goodness of persons, character traits, and other things, and such as the rightness or wrongness of actions, are natural properties of the same general sort as properties investigated by the sciences, and

² As Copp provides no lengthy discussion about constructivism, we will make use of Sharon Street's (2010) definitions, which the author uses to determine the type of constructivism being defended.

³ Cf. Papineau, 2015.

⁴ We use Jamie Lenman's (2006) definition of natural facts: “Natural facts are understood to be facts about the natural world, facts of the sort in which the natural sciences trade.”

⁵ Harman, 2012, 10.

(b) that they are to be investigated in the same general way that we investigate those properties.⁶

Nevertheless, this definition is too broad, for it depends on each theory to determine what is regarded as a natural property, what the investigation method is, and the definition of natural world. As previously mentioned, we must understand that this is just the *standard* definition of naturalism, for some of its further conceptions extend to constructivists, expressivists, relativists and even error theorists. Such is the range of naturalism over the families of meta-ethical views.⁷ However, it is ensured, as much as in *lato sensu* naturalism, the strong conviction in the rejection of dependence on any supernatural beliefs or commands.

Moreover, moral naturalism intends to demonstrate that moral questions can be explained from the perspective of a natural world, resulting in the onus of explaining why, when reducing value and moral knowledge to a purely scientific matter, it does not undermine the very condition of the moral world.⁸ Particularly in Copp's theory, the presupposition that moral facts and propositions exist – being such facts and properties reducible to natural facts and properties – points to a difficulty in establishing an association between naturalism and constructivism.

REDUCTIONIST AND NON-REDUCTIONIST NATURALISM

Copp follows the *standard* definition of naturalism. In other words, he supports a kind moral realism when he affirms that:

Moral *naturalism* holds that in thinking of things as morally right or wrong, good or bad, we ascribe moral properties to these things – properties such as moral rightness and wrongness, goodness and evil. It holds that there *are* such properties, and it adds that these properties are ordinary garden-variety natural properties – properties that have the same basic metaphysical and epistemological status as the properties a tree can have of being deciduous, and the property a piece of paper can have of being an Australian twenty-dollar bill.⁹

⁶ Sturgeon, 2006, 92.

⁷ An example is the relativist position, which speaks of moral relativism as given by the different social conventions established. And, since the conventions are assumed as social facts that presuppose natural facts, relativism could be qualified as a form of naturalism. But this seems to go against the eternal *convention-versus-nature debate*. See Sturgeon, 2006, 93. This is what seems to weaken Copp's realistic position or at least it does not make it a kind of moral dogmatism. This point will be developed in the text in order to overcome the apparent contradiction between realism and constructivism defined as a tenable position.

⁸ Divine command theorists question precisely this point, since the resulting determinism would preclude the conditions for freedom, will and responsibility of the moral agent. On this matter, see the theological voluntarism of Philip L. Quinn in Copp, 2006, 63-90.

⁹ Copp, 2007, 249.

In this regard, moral properties such as goodness, justice and virtue are natural properties. Therefore, ethical naturalism, even if being a kind of moral realism, has the particular condition of assuming natural properties and is, thus, distinguished from other kinds of realism. As we will see later, Copp's naturalist proposal is different from other kinds of realism as it regards moral properties as *natural* properties. For now, the naturalist challenge is to present and support a positive answer to the question of the existence of moral properties that come from natural properties, because it is not clear how those properties relate to each other.

Differently from science, wherein observation and explanation of natural properties are progressively more precise and complex, ethics seems to rest on obscure grounds when attempts are made to determine moral properties such as wrongness. Let us make a comparison. From a scientific point of view, we can precisely understand how a murder has occurred by precisely determining how the projectile of a fire gun passed through the victim's body, how the organs were affected and the consequent death. However, while criminal science regards the *causa mortis* as sufficient to determine the causer of this event, it is up to ethics to determine, as Hume said, "where is fact that here we call crime; point it out; determine the time of its existence"¹⁰ In other words, we could not assign the moral property of wrongness to the murder and, without it, it would not be possible to assign guilt to the offender. However, this challenge of explaining how moral properties are extracted from natural properties has a clear answer to the naturalist: there are no doubts that moral properties exist in the same way as natural properties do and, strategically, he argues that moral judgments about right and wrong are also assertions about the world. Therefore, certain moral properties exist in the world and we are able to identify them as well as science does with other kinds of natural properties.

The question of the relationship between moral properties and natural properties divides the naturalists. The *lato sensu* definition of naturalism does not require the existence of moral properties as a unique nature, neither as a *sui generis* one, but it only requires those properties to be connected with other properties. And it is this very requirement of connection between natural properties and moral properties that divides the theorists into non-reductionists and reductionists. The non-reductionist moral theorists, such as Brink, Sturgeon and Boyd,¹¹ proponents of what became known as Cornell realism, claim that moral properties are irreducible to natural properties, although they recognize that moral judgments work in consonance with explanations about the world. On the other hand, moral reductionism claims that "(...) the moral properties are properties that can be identified as being other, non-moral properties, where such identification involves us picking out exactly which other properties the moral properties are."¹² Summarizing, reductionists assume that

¹⁰ Hume, Appendix I, 6; see too Putnam, 2002, Chapter 01.

¹¹ According to Jackson (1998), Cornell realism has three characteristics: "(a) ethical properties are identical to descriptive properties; (b) relevant statements of identity are required *a posteriori* and (c) any analysis of the predicates and ethical judgments in descriptive terms is possible." (144)

¹² Kirchin, 2012, 49.

moral properties exist in the world, and this world is the natural one. Hence, it seems reasonable to assume that there is a link connecting the natural and the moral worlds.

Now, if we move the focus of our discussion about moral justification to the nature of moral properties, we can clearly see a problem that precedes this relationship between natural and moral properties. If we are willing to define the reductionist relationship between moral and natural properties, we need to understand what a *natural* property is.¹³ Copp claims that a natural property must be understood as an empirical property; in other words, “(...) a property is natural if and only if any synthetic proposition about its instantiation that can be known could only be known empirically.”¹⁴ Therefore, naturalism is defined in terms of what we can and cannot conceive from experience. Kirchin calls synthetical reductionist naturalism¹⁵ models similar to Copp’s, considering that it would be possible to reduce moral properties to natural properties in a specific way (in this case, to empirical properties). According to this view, we would be able to map the natural world or to discover the mappings available in this world, which would allow us to identify the usage of moral properties in the world through empirical investigation. If this is the case, there seems to be a subset of natural properties inside the natural set, as much as science has a subset of assessable natural properties.

II – COPP’S SOCIETY-CENTERED MORAL THEORY

NORMATIVE CONSTRICTION AND THE PICTURE OF MORALITY

To Copp, moral questions are *Janus-faced*, because every time we state a value judgment about certain questions, we tend to assume that our moral convictions are beliefs “similar” to beliefs about chairs, tables and books. In other words, we

¹³ Copp identifies four strategies to address this issue and solve the problem of Moore’s open question argument: (A) reductionist proposals, (B) ostensive definitions, (C) metaphysical definitions and (D) epistemic definitions. See Copp, 2007, 36-39.

¹⁴ Copp, 2007, 36.

¹⁵ Kirchin, 2012, 69. Kirchin still distinguishes the analytical materialistic reductionism, which affirms the reductionist possibility as the synthetic position, but he replaces mapping through non-empirical methods with a conceptual analysis. At this point, there would be a difficulty in establishing the link between moral and natural property, especially for the analytical reductionist. His strategy is to perform a network analysis, that is, relate moral concepts to other equally moral concepts and explain how such concepts are individually identified with natural concepts. For example, the concept of goodness is explained in a set of propositions wherein its use is relevant, demonstrating the role that goodness plays in this conceptual scheme in moral terms. The next step is to demonstrate, as did Moore and Hare, for instance, the concept of good as our central moral concept. In turn, we can remove all other evaluative and normative terms, explaining the good from natural properties that describe goodness. Thus, we could reinterpret how such natural properties perform the correct function because they play the role of goodness. With this, there is an anchor in our system from which we have a natural property that plays the role of goodness. Consequently, we reduced moral properties to natural properties.

intuitively take them as being true. At the same time, we see the other side of the coin when we consider the existence of moral convictions as a doubtful truth, insofar as we ordinarily tend to observe our moral convictions as different from our beliefs about chairs, tables and books. Generally, this dichotomy follows all of our metaethical questions.

Hence, the discussion about normativity involves the metaethical challenge of developing a theory about the nature of moral judgments, at the same time as we need to discuss the truth conditions of such judgments. Copp's central concern is to demonstrate the relationship between natural properties and moral properties, explaining how natural properties might be normative properties.¹⁶ The safest choice would be, then, to commit to realism through a naturalist model, since "(...) our moral beliefs ascribe moral characteristics to things, characteristics such as goodness and rightness, and that these characteristics are *natural* characteristics, relevantly similar to ordinary properties of things, such as meteorological or economic properties."¹⁷

At this point, Copp resorts to the so-called Society-Centered Moral Theory¹⁸ as an example of moral naturalism. According to Copp, every society needs a social moral code that ensures a balanced coexistence among individuals. After all, it is the social moral code that guides our relationships and allows cooperation among members of society.¹⁹ If we want to understand the dynamics in this picture of morality, we need to understand three basic interrelated features to determine the social moral code: we need to (i) identify the nature of morality (of moral judgments), (ii) determine its justification and (iii) discuss the role of society in this scheme and in the determination of the two first elements.

¹⁶ Furthermore, Copp assumes what he calls normative internalism, according to which if moral properties exist, they are necessarily normative, that is, a property can only count as moral property if it is normative. Therefore, the failure to explain the normative implies the failure to explain moral properties. Copp follows with a critique of motivational internalism, drawing attention to the need to explain the link between moral belief and motivation. According to him, despite defending a normative internalism model for moral judgment issues, his theory endorses a kind of motivational externalism, although he reaffirms the distinction between his naturalism and Cornell realism that also employs a kind of motivational externalism. Cornell realism's weak point, according to Copp, is not the problem of motivation, but the failure in the explanation of the normativity of moral properties.

¹⁷ Copp, 2007, 02.

¹⁸ Copp defines society as follows: "A society is a multi-generational temporally extended population of persons, embracing a relatively closed network of relationships of friendship, affection, kinship, and cooperation in reproduction, and limited by the widest boundary of a distinctive and salient system of instrumental interaction that facilitates pursuit of the necessities of life and the priorities of the group's culture." (Copp, 2007, 22)

¹⁹ According to Copp, morality is a useful set of social rules if we think in terms of its *function*. In this regard, basic moral propositions about right and wrong – for instance, *lying is wrong!* – will be true if and only if they work in the best way for their purpose of social organization. If we take the basic moral proposition that lying is wrong, socially we could establish the prohibition of lying. In these terms, morality is the useful set of social rules. In other words, moral beliefs are simply useful fictions.

In this framework, any attempt to harmonize the elements of this picture into a consistent theory would require us to explain the normative. This is the so-called *normative constrictio*n, which claims that “(...) an adequate metaethical theory must explain what the normativity of moral judgment consists in.”²⁰ This means that, before we discuss the moral weight of certain propositions, the properties of these propositions must be normative ones. This moral framework’s formula has two nuclear components. We need a cognitivist theory of normative language, that is, a standard-based theory (*SBT*) that offers a general model of truth conditions to normative propositions. In other words, a moral proposition will be true if and only if the rule or pattern to which it refers is justified.

Moreover, we lack a justification theory of moral codes and patterns, which Copp calls society-centered theory (*SCT*). Societies need moral social codes, and this means that we need certain patterns to be justified so that our moral assertions can be justified and regarded as true or false. This issue must now be investigated.

A standard-based theory does not provide moral justification of a moral code by simply deriving basic norms from a certain moral view. On the contrary, a moral code is justified if and only if a certain society is rational when choosing this social moral code.²¹ This feature endorses that moral judgments need to be justified by a cognitivist theory of normative language (*SBT*), which provides a general model of justification conditions, for it is in this discussion that the constructivist bias of Copp’s theory will be established.²² He points to a distinction between moral propositions such as “lying is wrong” and moral standards such as the standard that forbids lying and that may be expressed with the imperative “Do not lie!” In this regard, standards are the semantic content expressed by imperatives, although they (the standards) are not beliefs nor represent the world in any way. This is the cognitivist theory of normative language proposed by Copp:

The schema is intended to be applicable to laying out the truth conditions for any kind of normative proposition. The schema says that a (pure and basic) normative proposition of type K is true if and only if a corresponding standard of type K has the K-relevant “truth-grounding” status.²³

Starting with that presupposition, the philosopher proposes *SBT* as a scheme capable of explaining truth conditions of moral propositions in terms of a *status* that is relevantly correspondent to moral patterns. Consider the example of etiquette and the need for standards that support its rules such as “Don’t speak with your mouth full”

²⁰ Copp, 2007, 03.

²¹ This view is presented in *Morality, Normativity, and Society* (1995), but he seems to abandon it in his 2007 work, when he realizes that his theory is troubled by the task of defining rationality. Afterwards, this discussion is resumed in order to distinguish the naturalist and constructivist aspects of his theory.

²² Cf. Copp, 1995, Introduction.

²³ Copp, 2007, 14.

or “Do not interrupt other people’s conversations.” Such rules are only held if there is a correspondent (K-relevant) indicating that it is rude to speak with the mouth full or that it is not adequate to interrupt the conversation of others without being invited. The same would stand in the moral dimension, because if we apply the *SBT* procedure, we will discover that a moral proposition will be true if and only if there is a corresponding standard, which must have a morally relevant truth condition. And this condition of being morally relevant is central for us to understand what status a moral standard must have insofar as the moral proposition is true.

Hence, there must be a truth-grounding status that ensures the truth of moral corresponding proposition. In this interplay between standards and moral propositions, the *SBT* intends to ensure the truth of a proposition by testing its corresponding standard. Therefore, if the standard is justified, the corresponding moral proposition also is.²⁴

Social moral codes may be given by a society-centered theory, which, in spite of being necessary, may vary from society to society. With the notion of *SBT*, Copp has only pointed out to the need of justification of standards, but he has not indicated what such justification consists of. Now, *SCT* will provide the procedure by which we will be able to present a justification theory of moral codes and patterns.

According to the latter,

a basic moral proposition is true only if a corresponding moral standard is included in or implied by the moral code the currency of which in the relevant society would enable the society better to serve its basic needs than would the currency of other sets of rules and better than would be the case if no set of rules had currency in the society.²⁵

In this case, moral patterns need to have a certain truth-grounding status that is relevant. Such status is the capability of being morally authoritative, which means having strength to oblige us. We are interested in knowing if the *SCT* is able to explain the authoritative status of the morality of these codes, for it is its relevant condition for grounding moral truth.

There is a concern with identifying what is a normative assertion, since it has many variants, such as normative assertions about legal, moral, epistemological and etiquette matters. Besides, they differ from non-normative assertions like “the snow is white” or “my cat is fat,” since those do not present a corresponding standard from the normative perspective. Resorting to *SBT*, Copp defines that normative assertions point to certain standards. In the moral sphere, such standards hold an appropriate status that is correspondent to this sphere, as well as there are corresponding status to the legal sphere, epistemological sphere, etc. (remember the example of etiquette), what makes us infer that the nature of these status depends on the sphere of their application. In the moral sphere, the status of the assertions in this context is itself normative; hence, moral propositions will be true if and only if the corresponding

²⁴ See Copp, 2007, 15.

²⁵ Copp, 2007, 17. This society should ensure physical integrity, cooperative integrity (internal social harmony), and peaceful and cooperative relationships with neighboring societies.

moral standards are appropriately justified.²⁶ Therefore, there is an intrinsic normativity to each moral proposition, that is, they naturally oblige us in Copp's understanding. On the other hand, *SCT* is concerned with explaining the conditions under which moral patterns should be relevantly justified, which implies, in Copp's case, stating that a certain moral code is only justified to a society if this society rationally chooses and endorses such moral rules.

Thus, it is by the combination of *SBT* and *SCI* that Copp expects to justify his realist naturalism. As he himself says, such components could be combined distinctly when separated, that is, the social-centered theory could endorse a non-cognitivist position, as well as the standard-based theory could work to justify other supernaturalist or non-naturalist non-cognitivist models.²⁷ However, the society-centered moral theory is, in Copp's understanding, a powerful combination between standard-based theory and society-centered theory for solving the problems of moral epistemology and justification, respectively. Resorting to it, we would be capable of explaining the normative content of moral propositions and determining the nature of these propositions. As we could see, Copp's naturalism makes us relate moral and natural properties by an explanation of normativity. And, as we will see, *SBT* points to a cognitivist and realist condition of his theory, but it is especially in *SCT* that we find elements for a constructivist interpretation of his theory.

III – MORAL CONSTRUCTIVISM AND MORAL REALISM

MORAL CONSTRUCTIVISM

I will now identify the *lato sensu* versions of both constructivism and realism. The *lato sensu* moral constructivist claims that moral facts or truths are constructed, that is, moral norms are truths made from a set of beliefs, which are our evidences in ethics. A moral realist such as David Brink defines constructivism as follows:

Moral constructivism:

- (1) There are moral facts or truths, and
- (2) these facts or truths are constituted by the evidence for them.²⁸

Obviously, this definition is infected by Brink's realist grasp, but the core of the constructivist model, which states that values are constructed and not discovered, is preserved. However, when Copp discusses the problem of constructivism, he makes use of Sharon Street's²⁹ terminology to divide it into

²⁶ Copp, 1995, 03.

²⁷ Copp, 1995, 06.

²⁸ Brink, 1989, 20. Moreover, Brink claims that there are two kinds of constructivism in these terms: (a) relativist constructivism and (b) non-relativist constructivism.

²⁹ See his paper *What is constructivism in ethics and metaethics?* (2010).

(a) *restrictivist view*, according to which constructivism is a matter of substantive normative ethics, and (b) *metaethical view*, according to which constructivism has to deal with central second order questions. Street defines restrictivist constructivism as follows:

(a) Restricted constructivist views in ethics specify some *restricted* set of normative claims and say that the truth of a claim falling within that set consists in that claim's being entailed from within the practical point of view, where the practical point is given some *substantive* characterization.³⁰

Since it is a first order question, the restrictive proposal has the objective of accounting for the truth of a limited subset of normative claims. In this regard, a specific set of substantive normative claims chosen by the proponent of the theory, that is, the moral theorists, is talked about. Generally, this kind of constructivism claims that the truth of moral assertions comes from the fact that such assertions are forged from a predetermined practical standpoint,³¹ such as Copp's society-centered procedure or even John Rawls's original position. This view is endorsed by Darwall, Gibbard & Railton when they claim that this kind of constructivism deals only with first order problems and that, hence, it must explain, when facing different normative models, why it should be established instead of any of those other models.³² Summarizing, the constructivist view differs from other ethical views by prioritizing the relationship to be established between certain procedure and the correct/true result of its *use*.

Given its restrictivist character, metaethical questions still remain open in this constructivist view; hence, Street suggests the abandonment of procedural characterization and the acceptance of the practical standpoint as a determining one. In other words, she claims that metaethical constructivism is better, since the strength of the argumentation is not exactly due to the procedure, but to the point of view of the agent who endorses certain values, that is, to the practical standpoint about what to do and what not to do within this procedure.

Street proposes, then, the replacement of the procedural title of constructivism – represented by the slogan *no normative truth independent from the procedure* – by the notion of constructivism from a certain practical standpoint, here represented by the idea of *no normative truth independent from the practical standpoint*. Doing so, we are able to claim that there are no

³⁰ Street, 2010, 08. Similarly, Darwall, Railton and Gibbard (1992) corroborate this view: "(...) the constructivist is a hypothetical proceduralist. He endorses some hypothetical procedure as determining which principles constitute valid standards of morality. The procedure might be one of coming to agreement on a social contract, or it might be, say, one of deciding which moral code to support for one's society. A proceduralist, then, maintains there are no moral facts independent of the finding that a certain hypothetical procedure would have such and such an upshot." (140)

³¹ To Street, "(...) according to the proceduralist characterization, constructivist views understand normative truth as not merely *uncovered by* or *coinciding with* the outcome of a certain procedure, but as *constituted* by emergence from that procedure." (Street, 2010, 03)

³² See Toward Fin de siècle Ethics: Some Trends (1992), especially 137-144.

correction patterns within the normative realm, except from *some place*, that is, “(...) from the point of view of someone who already accepts some normative judgments or other – the point of view of a valuing creature.”³³ In other words, given that metaethical constructivism is based on the practical standpoint, these norms are not defined by the procedure, but by the agents who *create* the moral judgments.

Based on this distinction, Copp claims that his *SCT* – which he calls *original theory* since it was developed firstly (1995) – is a constructivist theory. However, later in the development of his theory he abandons this position, claiming that *SCT* – which he now calls a *basic theory* (2007) – can do well without these constructivist outlines. According to his understanding, the original theory is constructivist insofar as

(...) it is committed to the thesis that moral propositions are “made true” by a relevant kind of “endorsement” from a preferred “standpoint.” We can say that a constructivist theory defines an “endorsement function” that takes a specified kind of input and yields moral propositions (or moral propositions of a certain kind) as output; a constructivist theory holds that there is an endorsement function *F* such that a moral proposition (of kind *K*) is true just in case (and because) *F* yields the proposition as “output” given a relevant “input.” If we take this to be the central doctrine of constructivism, the original theory is constructivist.³⁴

On the other hand, *SCT*'s basic theory is not constructivist, because the truth of a moral proposition depends only on “(...) which system of norms could better meet the needs of the society in question, and this is an endorsement-independent matter.”³⁵ As it does not require endorsement, the basic theory would not need to presuppose a theory of practical rationality for the problem of justification. Thus, since it does not need rational endorsement, Copp's naturalist theory would be in the following situation: on the one hand, there would be original *SCT* as a kind of restrictivist constructivism insofar as it requires a procedure in which rational agents endorse and, consequently, construct the *social moral code*; on the other hand, basic *SCT* would not make use of a theory of rationality and would not reduce the problem of morality's normativity to a simple matter of rational choice theory.³⁶ Despite the abandonment of original theory, Copp does not explain exactly why his theory is no longer constructivist, for *SCT* seems to be a restrictivist kind of constructivism and, although it might not settle the core of justification on the rationality of the involved agents, it still has to explain the transition of natural properties (input) to moral properties (output). Perhaps this explanation might be given by the realist definition of his theory. Let us see.

³³ Street, 2010, 05.

³⁴ Copp, 2007, 19.

³⁵ Copp, 2007, 20.

³⁶ See Copp, 2007, 21.

MORAL REALISM

In his work *Moral Realism and the Foundations of Ethics* (1989), David Brink provides a broad definition of moral realism.

Moral realism:

- (1) There are moral facts or truths, and
- (2) these facts or truths are independent from the evidence for them.³⁷

Together, (1) and (2) constitute the hard core of *lato sensu* realist theories. In terms of moral ontology, this definition is really broad, to a point that antirealists³⁸ could agree with (1) and the realists themselves would diverge – as they currently do – about *where* to find (2). In order to clarify this point, Copp takes premises (1) and (2) as a starting point, but points out to five characteristics of realism, distinguishing it into basic realism and stance-independent realism:

Basic realism

1. There are moral properties (and relations). There is, for example, such a thing as wrongness.
2. Some moral properties are instantiated. For example, some actions are wrong.
3. Moral predicates are used to ascribe moral properties. When we call an action “wrong,” we are ascribing to it the property wrongness.
4. Moral assertions express moral beliefs. When we call an action “wrong,” we are expressing the belief that the action is wrong.

³⁷ Brink, 1989, 16

³⁸ The classic example of this dissolution of the realist core is John Mackie’s antirealism (1977), although his critique seems to deal with only certain kinds of realism. His position is clearly antirealist insofar as it endorses three basic theses of moral ontology: a. There are no objective moral values (moral facts) that are independent from the subject’s mind; b. Moral judgments are independent from objective moral values (if they exist); c. Objective moral values are derived from moral judgments. See Mackie, 1977, Chapter 01. As a result, his proposal takes the form of an *error theory* to inform the mistake of assuming that setting out moral judgments would necessarily imply presupposing objective moral properties. This assumption is a mistake. Therefore, Mackie calls for the famous arguments *from relativity* and *from queerness*. The former provides indirect support to second-order skepticism to confirm the existence of discordant and antithetical moral judgments, which are derived from their respective contexts and time. The latter is divided into two parts: the metaphysical element, which states that “(...) if there were objective values, then they would be entities or qualities or relations of a very strange sort, utterly different from anything else in the universe” (38). In the same measure, the epistemological element states that “(...) if we were aware of them, it would have to be by some special faculty of moral perception or intuition, utterly different from our ordinary ways of knowing everything else.” (38)

5. Moral properties, in being properties, have the metaphysical status that any other property has, whatever that status is.³⁹

Stance-independent realism adds a sixth (6)⁴⁰ thesis, claiming that certain moral facts exist in a mind-independent way. Such problem of moral ontology interests us in the discussion about naturalism, for the possible constructivist aspect of Copp's theory would be supported on the basis of moral realism *and* constructivism. We have established, thus, the broad and the strict definitions of constructivism and realism, remaining, now, the discussion about how the relationship between them is established in a naturalist manner.

In a nutshell, the philosopher argues that the metaethical constructivist might accept the five-theses package of basic realism, but would fall into contradiction if he accepted thesis 6 (stance-independent). Concerning thesis (1), about the existence of moral properties, we can say that the *lato sensu* constructivist would endorse the existence of such properties, although he would not reduce them to natural properties, nor search for a value-independent order for them. In this regard, the constructivist's adherence to this thesis does not make him a reductionist naturalist. As to thesis (2), the constructivist can also accept that there is a relationship between moral and natural properties, just like in the naturalist case, but with the clarification that non-moral properties, since the moment they take part in the constructivist procedure, are interpreted as moral properties, which does not seem to be a problem for the naturalist.⁴¹ About thesis (3), the constructivist has no problem in assuming the existence of instantiated properties, inasmuch as they are the result of a construction process, that is, created by its evidence. Regarding thesis (4), the constructivist – specially the one with a restrictivist bias – claims that the predicates are ordinary only until they take part in the procedure. From this point on, they would become moral predicates, non-reducible to simple scientific verification. The example presented by Copp is claiming that the predicate *wrongness* is used to assign the moral property *is wrong* to action x, whereas such predicate would not hold any *sui generis* condition for its existence. This discussion is directly connected to thesis (5), since constructivism would not have to deny that the agent's mental state reduces normative assertions to ordinary beliefs, insofar as the agent's mental state is of valuation and distinguishes moral assertions from simple ordinary beliefs. In this regard, beliefs would not be reduced only to its truth conditions, for the constructivist

³⁹ Copp, 2014, 121-122. According to Copp, non-cognitivist antirealists could agree with thesis (1), (2), (3) and (4), but point (5) would be paramount to distinguish his position. Thesis (5), that moral properties share the same metaphysical status of any other property – regardless of the status – would distinguish Copp's position from deflated versions of antirealism and non-cognitivism.

⁴⁰ In his text *Morality in a Natural World* (2007), Copp provides another definition of the sixth thesis: "(6) Moral properties are natural properties." (10)

⁴¹ Examples of this defense are the reductionist naturalism sustained by Copp and Peter Railton and some trends of neo-Aristotelianism and post-positivism. Cf. Darwall, Gibbard & Railton, 1992, 11.

would regard them as evidences only in the first moment, that is, before they join the *procedural game*.⁴² Regarding the example above, constructivism would agree with the use of the predicate *wrong* in propositions such as “Torturing is wrong!” By doing so, we would assign the moral property *wrongness* to torture, although this property preserves its condition of being a belief that holds epistemic value.

THESIS (6) – THE PROBLEM OF MIND-INDEPENDENT MORAL REALISM

As we have seen, the realist theses from (1) to (5) of the so-called basic realism might be endorsed by the constructivist, but what is there to say about thesis (6), which defends a value-independent order and a kind of independent moral realism? Undoubtedly, the big restriction to a relationship between realism and constructivism in Copp’s naturalist theory concerns this sixth thesis. In a broad spectrum, independent moral realism (*Thesis IMR*)⁴³ is a kind of ontological defense over the existence of an independent moral fact or truth, that is, regardless of the mental states of the moral agents. This strong type of moral realism, in which a certain moral fact or truth seems to be resistant to possible *distortions* given by the agent’s interpretation, would undermine the possibility of a moral constructivism that admits that moral facts or truths exist and are constructed in spite of any individual or social group. Hence, the mind-(in)dependence problem would distinguish realism supporters from *lato sensu* constructivism supporters:

Realism:

- (1) There are facts or truths of kind x, or
- (2) These facts or truths are independent of the evidence for them. Constructivism affirms R(1) and denies R(2).

Constructivism:

- (1) There are facts or truths of kind x, and
- (2) These facts or truths are constituted by the evidence for them.

Constructivism asserts and realism denies that the facts or truths in question are constituted by our evidence for them.⁴⁴

Briefly, we could say that facts are evidence-independent and moral truths are evidence-independent to supporters of thesis (6). But, what exactly is this

⁴² Hare draws attention to ethical descriptivism, which has incurred in what Austin called “descriptive fallacy”. See Hare, 2003, 71-94.

⁴³ Kirchin defines this position as follows: “*Thesis IMR*: The existence of moral properties and moral reasons is a mind-independent matter. That is, the existence of such properties and reasons is not dependent on what human beings, either individually or collectively, think, desire, are committed to, wish for, etc. Similarly, the ‘type’ or ‘character’ of value or reason that they are - e.g. goodness, kindness- is a mind-independent matter.” (2012, 22)

⁴⁴ Brink, 1989, 16.

independence? Brink suggests three kinds of independence: causal independence, metaphysical independence and conceptual independence.⁴⁵ Causal independence would be supported by scientific realism models when claiming that objects like doors and walls exist as causally independent of their creator's mental state. As to the moral realist, his concern is with a kind of *conceptual* independence (such as Moore trying to isolate the Good from our definitions and frameworks), or even with *metaphysical* independence as held by Platonist moral intuitionism. In this regard, conceptual or metaphysical independence works on grounds of a fact-value dichotomy. Summarizing, these theories claim that properties of moral terms and their truth conditions are not provided by the moral agent, which means they constitute this value-independent order.

On the other hand, constructivist theories support that facts are evidence-dependent. Hence, moral truths are evidence-dependent, as well as moral judgments are mind-dependent. Mind-dependent models are models in which truth conditions of moral judgments evoke counterfactual assertions about the agent's choices in hypothetical circumstances of some kind. As said before, constructivism is a kind of mind-dependent defense when claiming that the truth of moral judgments depend on knowing if they (the judgments) are coherently related to the moral principles endorsed by valuing creatures engaged in an idealized reasoning process.⁴⁶ If this is the case, thesis (6) cannot be claimed by the constructivist, because it would result in a contradiction. As mentioned, if Copp assumes the strong type of realism (endorsing thesis 6), according to which moral reality is fixed regardless of any of the agent's beliefs related to it, the application of constructivism would be redundant.

A remaining question is: how does Copp face this dilemma in his theory? As seen before, he identifies in the moral landscape the need for normative judgments to have strong truth conditions, as well as the need for the determination of those conditions. Moreover, it seems that we must determine, at some point, a kind of *normative leap* from natural properties to moral properties, which inclines us to think of constructivism as an adequate methodology. We can assume restrictivist constructivism and its procedural use, which defines the conditions that will produce moral facts or truths as a result (*output*). Hence, such facts or truths, and only those, will be *moral* facts or truths, due to being generated by the procedure. In other words, the procedure provides the conditions of justification of moral propositions, although it does not explain the resulting kind of fact or truth in terms of moral objectivity. What we know is that constructivism claims that moral facts are determined as the result of a procedure made on the basis of a value-dependent order; thus, it would accept basic realism without contradiction, that is, it would endorse theses (1) to (5) pointed out by Copp. That would be a procedural interpretation of the original theory of *SCT*, to which some judgments would be justified as social moral codes.

⁴⁵ See Brink, 1989, 15.

⁴⁶ See Copp, 2013, 115.

Also, if we assume Copp's model with a kind of metaethical constructivism – considering the mind-dependence question to the normative question and the consequent view to which there is no value-independent order to the practical standpoint –, his theory would still endorse a counterfactual dependence of value on the attitudes of agents. To both types of constructivism (procedural and metaethical), the essential consideration is that there are no moral facts beyond the standpoints of the agents who value from the society-centered perspective. Therefore, every social moral rule would be created and justified from this point; before that, natural properties would not hold the sufficient normative weight to be a justified moral property. *Per se*, *SBT* only provides the truth conditions of normative propositions as a cognitivist theory of normative language, which requires the condition of *SBT* to provide a kind of justification to social moral rules. Now it seems to point to a *moral construction*.

CONCLUSION

With these explanations, I believe we have enough elements to consider the relationship between constructivism and naturalism in David Copp from a common basis: moral realism. As mentioned before, Copp's moral naturalism adds a sixth thesis to basic realism: (6) moral propositions are natural propositions. If Copp assumes his realist naturalism to be consonant with constructivism, the thesis above will lead him into contradiction. If he assumes similarities between naturalism and constructivism, then his realism cannot resort to the present mind-independence thesis (6). By doing so, he would render constructivism irrelevant to the justification perspective of his theory. However, Copp's naturalist proposal points out that moral properties might be understood as natural properties. In this case, it seems that the constructivist is able to endorse the naturalist position of a basic realism, as far as the *construction* determines what moral facts or truths are.

The fact is that resorting to *Society-Centered Theory* seems to point to some kind of constructivist naturalism in Copp's proposal, for if we assume a rationally required justification for moral rules, 'correct' and 'incorrect' are not determined beforehand. In other words, the ideal moral code would be the moral code rationally chosen by a society to be used as a social moral code, regarding the needs and values of this society, as well as its circumstances. If all the members of a community subscribe to it, then the ideal moral code determines what actions might be considered wrong or right in relation to this society, which seems to reveal a relativist character in his theory.

Now, if the constructivist option presented above is not sufficient, Copp can still assume metaethical constructivism as being compatible with his theory's realism, accepting even his five theses on realism. However, the shadow of the sixth thesis still hovers over his theory. In this case, metaethical constructivism claims that normative facts are constructed by facts derived from rules of practical reason, which are combined to non-normative facts. The trick would be to provide a plausible model of practical reason, which implies explaining what involves *playing the game of valuing*.

The aim of this short draft was not to support Copp's constructivist claims, but to clarify some obscure points about this theme throughout his naturalist theory, since the relationship between naturalism and realism is not a cause of disagreement. If Copp is successful in his normative *desideratum* of integrating natural basis to a constructivist model, it remains to be seen to what extent the commitment to the natural world negatively affects the mind-dependence condition of constructivism, since the naturalization of morals implies that normative explanations are derived from scientific explanations. Extrapolating this relationship, from the perspective of having normative reasons for actions, the fact of knowing scientific data about what we regard as a *good life* does not seem to be a relevant one. Reductions on the answer about the social moral code make the facts about what is valuable dependent on the result of causal process, which might be irrelevant to moral justification. It is due to the fact that, as in the previous example of murder, the moral condition of an action does not include the scientific explanation of the rule based on the victim's *causa mortis*. In this regard, we notice that the reduction of moral properties to natural properties is derived from the answer of agents under idealized circumstances, such as *SCT*. In this case, the question is about normativity and not about moral itself. When the focus of the discussion is moved and the central question is obscured, we take the risk of reducing not only moral properties to natural properties, but also the moral realm to the normative realm. In this case, we would no longer be asking ourselves whether moral rules, such as chess or baseball rules, are *good rules* but, instead, we would be asking why we should observe them as we play. Even though antirealist criticism has been refuted in a minimum level, assuming a naturalized kind of morals does not avoid old embarrassments, insofar as presenting a set of factual data to explain the value of a certain moral judgment does not seem to be enough to *demystify* the nature of moral judgments. Such is the problem of not assuming the constructivist condition of Copp's theory. Without it, the *mystery* remains to be seen...

IS A PHENOMENOLOGICAL PERSPECTIVE ON THE SOCIAL CONSTRUCTION OF REALITY STILL VIABLE?

FABRÍCIO PONTIN

What I want to offer in this article is a reading of political philosophy, particularly modern and liberal accounts of political philosophy, within a phenomenological perspective. My hypothesis is that a re-articulation of the main issues of phenomenology as it relates to social philosophy, first in the context of a natural-world, non-transcendental, description (as developed by Schutz), but also in the context of a transcendental, in-depth, approach (as developed by Waldenfels), provides us with a methodological ground that can respond to some of the issues raised in the Habermas-Rawls debate, adding some important insights for many of the paradoxes identified in the praxis of Democracy. It will also provide us with clues as to how to think issues of social choice and decision making outside the framework of a methodological individualism.

Weber had divided his narrative of social action into three levels to illustrate how we pass from a level of individual conceptualization of judgments and creation of identity (I), into an attempt to communicate these concepts and “mean to say” something to some other person (II), until finally one is also hit with a somewhat different conceptualization that motivate social interaction (III). Habermas, on his turn, will stress the first aspect of this process of conceptualization, claiming that sentences can and will be understood if the parties operate rationally. Moreover, not only will conceptualizations be understood, they will be agreed upon. So the universal of discursivity, which is understood aprioristically, is articulated with a universal will for communication.

Schutz, however, tries to point at the static narrative behind these levels of understanding and action, providing a different account for the Weberian analysis of meaning and its articulation with action. Schutz identifies that Weber’s preoccupations, at the end of the 19th century, are not very different to those of his contemporary in philosophy, Edmund Husserl. To a great extent, the analysis of meaning constitution in Weber is almost identical to what we find in the Logical Investigations. In the idea of the solitary life of the soul, for example, Husserl also provides us with a first analysis of identity formation as “solitary” and isolated, and then moves such formation in terms of expression and indication, which are modes in which pre-conceptions about the world will enter interaction. Schutz points at these similarities in a brilliant way, but he also holds that Weber’s static understanding of action fails to describe social action and interaction in terms that are fair to experience. The problem of meaning

constitution, identity, and expression, in Schutz, is a time-problem.¹ This means that all processes that lead to the formation of an identity and how an individual acts in society will be understood in terms of time-consciousness.

This genetic level of analysis is understood by Habermas in terms of a philosophy of consciousness which is operative on the “lifeworld paradigm”. This motivates Habermas to see system-theory, as developed in the social sciences, as an heir to the reflections in Schutz and Husserl.² Habermas, however, insists on the possibility of consensus as a guiding principle for sociological reflection: social interaction presupposes that individuals are in consensus about the meaning-like structures of the world and the value-like expressions of language. This element in Habermas is in tension with the account given by Rawls and with a phenomenological account of social action. Such tension reproduces the before-mentioned epistemological aspect of political liberalism: the question of language and discourse as a leading clue to the constitution of society and politics.

Schutz points at a less formal understanding of the process of social-constitution, and stresses the importance of historicizing the elements of action that characterize the behavior of individuals in society. In this sense, Schutz already anticipates a critical element in the Habermasian development of Communicative Action, that is: stressing the element of consensus in social interaction is to introduce an *ought-like* element in sociological analysis. Now, if we wish to understand the paradoxes in social action, especially as they relate to advanced democracies and their peculiar character, we need to move away from a description of how language and communicative practices *should* operate in society, and investigate *how* they are given in social interactions. To be sure, Habermas provides us with a sophisticated description of the processes of social interaction that lead to consensus in deliberative practices. The problem is that such description is artificial.

This is not to say that Schutz himself does not have limitations. His genetic analysis of the processes of social constitution still trusts a somewhat formalistic account of the lifeworld. Schutz is successful in showing how the process of action and world-constitution can be accounted for phenomenologically, but his form of description of the “rational choice of values” in society is limited by his understanding of the lifeworld in terms of time-consciousness. This means that Schutz is not so open to the ways in which our constitution of values is determined by the environment in which these constitutions appear. Schutz still

¹ Alfred Schutz. *The Phenomenology of the Social World*, (Evanston:Northwestern University Press,1967), 12

² “There is no doubt that the familiar psychological and sociological models of an isolated actor in a situation, affected by stimuli or acting according to plans, gain a certain depth of focus through being connected with phenomenological analysis of lifeworlds and action situation. And this is in turn the jumping-off point for a phenomenologically informed systems theory. This shows, incidentally, how easy it is for systems theory to become the heir to the philosophy of consciousness.” Jurgen Habermas. *The Theory of Communicative Action, Vol. II*. (Boston: Beacon Press, 1987), 128.

understands the emergence of ideal types in a Weberian way, that is, types are ultimately related to the history of an individual that articulate his optimum notions of values. But Husserl developed his idea of types in a generative level later in his philosophy, and this brings consequences to the form in which we understand society.

Such comprehension of the processes of value-constitution, which also constitutes the political sphere, will allow us to understand some of the main issues in contemporary politics in a different light. Particularly, it motivates us to look into the evolution of the argument for a relationship of meaning and action, particularly in Kant, Weber and Husserl. The questions of voting, participation and what it means to have a preference, for example, will have to be reconsidered under these lights, and it is my thesis that a phenomenological analysis of these issues will allow us to comprehend the crossroads in political liberalism in terms of a permanent crisis in the processes of signification and comprehension of world-views, one that can be understood as both destructive (in terms of political totalitarianism) and emancipatory (in terms of development of personal identity and plans of life).

Ultimately, it is my take that generative phenomenology is a phenomenology of human freedom and that the legacy of a phenomenological interpretation of political philosophy points at a defense of freedom of speech. Husserl's generative phenomenology, in fact, points at the absurd of the totalitarian descriptions of human nature and vocation, and the static, stratified, notion of society that we find in contemporary philosophers like Heidegger and Scheler. It is particularly important to focus on this aspect, given the dark political implications of Heidegger's philosophy. Phenomenologists, especially social phenomenologists, must admit that Heidegger's political options, in the late 1920's and early 1930's are a direct consequence of his epistemic choices and options, particularly his understanding of territory, ground, identity and vocation. There is an important collaboration in taking a Husserlian point of view here; as it shows that not all phenomenological contributions to political and social philosophy will lead to an anti-humanistic, anti-modern, interpretation of personality and rights. Husserl places himself as an heir to the philosophy of Kant, Hegel and Descartes while Heidegger points at the *Jüdische Sprachen* that threatens the motherland, the German territory, and identity. As I approach Tugendhat's critique of Heidegger's epistemological options, this contrast will become central to my option of dropping a non-liberal interpretation of political philosophy within the phenomenological spectrum.

I hope to be able to show, at the end of this article, that a phenomenological approach to political philosophy will enrich our understanding of the history of human rights. Moreover, it will allow us to pin-point at some of the apparent paradoxes of political philosophy as a hermeneutic problem: we identify as paradoxes, as irrational choices, choices that we are not able to situate historically. Understanding these historical circumstances wherein different perspectives are given to social constitution will help us think through some of the political issues we struggle with nowadays, particularly the questions of social identity, and the current tension between different cultures and

implementation of rights and culture claims— which has been the main difficulty liberal societies have been struggling against.

NORMAL AND OPTIMAL / HOME AND ALIEN: ELEMENTS FOR A SOCIAL COMPREHENSION OF THE LIFEWORLD

The question of the constitution of the normal and the optimal, in Husserl, plays a similar role to the one we attributed to the notion of type in Weber. The normal refers to a process of habituation: in time, individuals incorporate elements of their surroundings, and the set of experiences they have in these surroundings, to their consciousness. Such processes create familiar appearances and relations of expectancies and disappointments. Allow me to give a banal example here in the form etiquette norms: if one grows up in a scenario where if one sneezes, one immediately apologizes for it, this creates a norm. This norm, then, orients comportment in every point in which one sneezes. In that case, if one fails to apologize after one sneezes, other actors who share the same heritage will react with disappointment.

Now, though this example is banal, it points at certain characteristics of the constitution of normality. First of all, this stresses that we will only be able to understand something as a norm as it is inserted in time. I already anticipated this while bringing to modalities of time-consciousness (expectancy and disappointment) in my description of normality. It is because of this timely constitution of the normal that I am able to establish a concordance in the continuity of appearances, leading to my qualification of a certain experience as “normal”³.

As these normal experiences are repeated in time, they acquire a certain density, and we might then qualify our normal experiences as “optimal”. An “optimal” here does not imply a rupture in the former experiences of the “normal”, at least not yet. It seems to me that an optimal level of constitution is at first possible in terms of an experiential peak for a certain normal, say, in terms of etiquette, participating in a tea-ceremony in Japan might institute an optimum in terms of table manners, this optimum does not necessarily introduce a rupture, or a discordance, with my previous set of experiences regarding table manners, but it does introduce a standard against which all other experiences of the same type will be compared.

But a rupture in this experience, a discordance with the way the object normally appears, that is an abnormality, can re-constitute the object. The abnormal here will have its own continuity, its own presence in the experience of the object. A discordance will then constitute, an optimum, a new norm. This emerges out of the experience of the everyday world, of the lived-world. Of course, this is fundamental to our understanding of home and alien worlds; the normal and the abnormal only apply in a system of specific species, in a determined ecosystem, if you wish.

The interesting point here is that a discordance is a break within a former

³ Ibid, 132

order of normally instituted constitution. If in the example of the tea-ceremony we have a confirmation of previous established norms of table-manners, but on a higher level of technical and aesthetical proficiency, in this case we would have a completely different dimension of what table-manners mean to begin with - in this case, the abnormal institutes a new level of signification, and points at a different mode of constituting social reality.

In fact, abnormal forms of constitution reflect a multiplicity in the way in which we constitute social reality. This multiplicity, in Husserl, is expressed in the distinction between a familiar homeworld and a strange alienworld.

Now, the familiar and the strange are generated, in Husserl, according to habituation and experience. In this, I am dangerously close to simply repeating the distinction between primary and secondary socialization in Weber.

A homeworld would operate in a very similar form to the primary socialization in Weber, as it is the immediate, familiar, surroundings in which individuals appropriate and institute norms in accordance to a shared heritage, a shared socialization, a shared set of meaning-complexes. Husserl, however, stresses the importance of presentational and passively synthesized information in the constitution of this “home”. In that sense, the formation of a homeworld is dependent both on conscious activity of a self towards the constitution of meaning-like relations to objects that surround oneself, as well as non-conscious affection of that self by previously intersubjectively constructed and instituted relations, and contingent and territorial circumstances. In that sense, the order of an *oikonomia* is established historically within a space of similitude, and within this space individuals will also develop more and less dense experiences that will be further qualified as “normal” or “optimal”.

Now, an alienworld takes the place of secondary socialization: an alien form of worldliness is one that disrupts the pacified set of experiences that we had previously constituted at home. At this stage, we have a reconsideration of the fundamentals of a familiarly constituted norm, now in relation to the emergence of a discordance, of a realization that something can be constituted in a different form. These processes of cultural shock insert an element of disorder into a previously well-ordered constitution of an object. Such disorder is, above all, temporal: we cannot grasp where the disordered element “fits” within our former experiences regarding a type – and this brings the realization that the temporal constitution of norms is not shared universally.

Now, is this to say that the transcendental aspect of the lifeworld is in jeopardy? After all, how can we attest to the structure of the lifeworld if the historical conceptualizations of objects differ so radically?

It seems to me that Husserl would indicate that to ask this question would miss the point of the transcendental analysis of the lifeworld. But I also believe that Schutz identified at this point a contradiction in the analysis of the lifeworld. I will get back at this in detail further in this article, but for now it suffices to say that Schutz, perhaps for lack of access, or for a general misunderstanding of the scope of the analysis of the lifeworld, does not comprehend how the transcendental character of the lifeworld refers to the co-generativity of home and alien. This is to say that the generation of normal and abnormal types is

always being constituted within an universally shared lifeworld which shapes the possibilities for the constitution of society. Hence the distinction between an analysis of the *constitution* of social reality, as in Weber, and an analysis of the *possibility* of social reality, which is what we have in Husserl.

Further, I will deal with how Schutz uses Husserl to approach and criticize Weber, but I also want to stress the limitations of the approach taken by Schutz, and how his decision to focus on a formal analysis of the lifeworld might have opened the way for Habermas' critique of interpretative sociology, and his adoption of Tugendhat's critique of Husserl's phenomenology of consciousness.

At this point in the dissertation I have enough elements to move to a phenomenological contribution to the issues of political and social philosophy. Schutz was able to identify in Weber a great affinity with the main problems in Husserl's phenomenology. Weber and Husserl share a general assumption that the objective world, without individuals to experience it, is meaningless, and that the structure of "empirical reality" or "objective reality" is constituted in history. Now, we will see that Husserl was able to move further than Weber in this issue, particularly in the question of historicity and the division of formal and transcendental approaches to knowledge. This division is important to the understanding of the field and scope of social investigation, as well as to the epistemic consequences of taking a transcendental perspective to political and social issues.

Next, I will attempt to formulate how Schutz relies on Husserl to develop a non-transcendental perspective to an analysis of everyday life⁴, in doing so, I hope to clarify the main differences in the historical and contextual account of meaning, as provided by Weber, and the phenomenological account given first by Husserl, and then by Schutz.

Schutz appropriates the entire *lexicon* of Weberian sociology, infiltrating some elements of phenomenology for two effects. First, Schutz wants to broaden the scope of Weber's sociology, by reframing the question of the "terms" in which we argue the structure of social reality. Schutz, in this sense, anticipates a problem in Habermas' normative view of society: the structure of social reality cannot be just a matter of intersubjective agreement to what the terms used to describe external objects should be⁵, rather, the question of sense and meaning is left unsolved in classical sociology precisely because it is taken for granted as a matter of raw epistemic collection of "world views". Second, Schutz informs the Weberian analysis of types with a structural phenomenology, which in turn allows him to disclose the structure of the social world as a structure of intelligible *intentional* meaning. Intentionality, of course, is the key turning point here: Weber speaks of "intended meaning" in an empirical sense, as in "what was meant" in a sentence uttered by a social actor; in Schutz intentional meaning refers to the structure of intentionality, trying to point that the social

⁴ Dan Zahavi, *Husserl and Transcendental Intersubjectivity*. (1996 repr. Athens:Ohio University Press, 2001), 22-24

⁵ Alfred Schutz, *The Phenomenology of the Social World*. Translated by George Walsh and Frederick Lehnert. (1932, repr. Chicago: Northwestern University Press, 1967), 9

behavior of an actor is already given in a context, and that the “individual” actor is also inserted in a social space. In that sense, Schutz dislocates the solipsistic development of solitary constitution of meaning in Weber, introducing a field wherein a social phenomenology could operate.

This dislocation operates on Schutz’s understanding of temporality and how it shapes behavior and reason – that is, how time-consciousness allows us to speak of the constitution of personality differently. Identity formation, in Weber, presupposes the existence of a primary form of socialization that informs and binds individual conceptions. Socialization, then, takes place within a certain social structure, and the process of formation of identity is immediately informed by social reality – by the surroundings of a social actor. This first movement is further made more complex in secondary form of socialization, but in this context the individual already has something like a “formed” personal identity. Thus, personal identity in Weber is dependent on the immediate aspect of socialization, on the form of early socialization of an individual; still, this dependency does not provide social actors with a static form of identity, further socialization might render aspects of primary socialization undesirable, or put them under a different light.

But the dynamic nature of identity constitution also re-constitutes the environment, or the social-organism, where actors are situated, so the social surroundings that once affected a social actor is now affected by one’s acts. This temporal dialectic of being-affected by social reality and constructing new conceptions onto social reality characterizes the social construction of the lifeworld. Such process is underdeveloped by Weber.

A social phenomenology will therefore presuppose Weber’s notion of meaning and social action, but also indicate that what Weber has offered is only at the level of a *regional* analysis of meaning and a *historical* understanding of temporality. This might appear of little consequence for a sociological analysis of everyday life, but it brings consequences to the way in which we understand decisions and judgments, and particularly how we attribute validity to beliefs about conduct in an increasingly diverse homeworld.

So we can follow the consequences of the formal approach taken by Schutz, I will break down his critique of Weberian sociology into three different topics: first, I will introduce the distinction between intended and observed meaning in Schutz as a re-articulation and critique of the distinction between absolute and instrumental understanding in Weber. Second, I will point to the way in which Schutz takes on the prevalence of perception of reality over the representation of reality, which allows him to read the notion of ideal types as at *loss* with experience rather than enriching it. Finally, I will point that social relationships, and social action, in Schutz are not meaning-oriented, but rather other-oriented.

THE CONSTITUTION OF MEANING IN SOCIAL PHENOMENOLOGY

Schutz defines meaning as a way in which we approach former experiences; this approach will circumscribe certain aspects of external reality as possessing a determined density, a materiality. Materiality is the building of a mastery of the

I, the realization that it is possible to have a perspective on an experience, and that this perspective, though it presupposes socialization, is a solitary endeavor. Only the individual that perceives something as meaningful is fully aware as to why and how that particular experience becomes charged with meaning. Later, one might wish to share his view, but as we will see communicating the entire density, the entire meaning, of a personal experience is always at loss with the actual experience. In this sense, having a self presupposes constituting what Schutz calls a “world of experiences”, that is, a scheme of references and presentations to which a social actor refers to in the moment something is constituted as meaningful. Further a self acquires a density, “a certain way of directing one’s gaze at an item of one’s own experience.”⁶ In this we create a peculiar stock of meaningful-like relations to objects and experiences, based on the repetition of these objects in experience.

One could be tempted to ask how this description is any different than the Weberian description of identity constitution. The main distinction is that Weber presupposes that meaning is always rationally constituted, that is, that it can be described in a detailed analysis of a conscious individual that is in full control of which elements of her primordial world of experiences are affecting her experiences. Her judgment of something as meaningful is perfectly coherent within a system of references, and behavioral factors do not enter play as “meaningful”.

In Schutz, references that are not represented rationally or consciously by an individual are relevant and participate in the constitution of meaning. Individuals never really cease to be affected by previous perceptions that inform and constrain the circle of what is meaningful. In that sense meaning is both *intentional*, as it refers to the form of conscious apprehension and signification of presentations, and *observational*, since all data which is presented to a social actor becomes relevant to the process of signification.

At this point, meaning is constituted topologically: it is attached to the place of speech and to the circumstances of a given place of speech. It also reflects the multiplicity of presentations as always already integrated within a context of meaning. *All* lived experiences, in Schutz, are meaningful insofar they constitute the field of possibilities in which something is perceived as meaningful⁷.

I want to focus on the question of *disorder* in the constitution of meaning within Schutz. I do not think it is controversial at all to claim that social action in Weber is a matter of ordering a previous stock of shared and individual knowledge and expressing this knowledge as a preference, as something that one values. In that sense, when I say that I choose *x* rather than *y*, I am expressing a historically constituted scheme of representations which is consciously chosen by me. I am expressing an ordered and rational interest. In some scenarios, if we are dealing with goods or values that are of importance to what I consider to be the core of my self, then I will attribute a static value to that good or value, that appears to me as absolute. In both cases, my preference is ordered, rational and fully under my control.

⁶ Ibid, 42

⁷ Ibid, 27

Such is not the case in Schutz. First of all, the elements that constitute something for me as meaningful are not fully well-ordered in my consciousness. There are affective elements in the process of object and value constitution that I simply cannot grasp, and yet, they participate in my attribution of “meaning” to that same object. Object constitution in Schutz is, in that sense, disordered.

Things get further complicated, then, when we think of terms of a social act and a social action. If these modes of social interaction, in Weber, reproduced the modes of rationally constituted meaning (as either instrumental or absolute, according to the orientation of a self towards that object or value), in Schutz an act includes all active and passive elements of the lived-experience of a social actor: instinctive behavior, for example, can and will play a part in the constitution of a social act. Action is not confined to the realm of the rational⁸.

In that sense, the notion of rational choice is taken under a very suspicious light in Schutz’ phenomenology. For Schutz, reducing the process of choice to a ranking of preferences that can be described in a transitive and connected form loses sight that any constitute preference is always in interaction with an individual who never ceases the constituting process. In that sense, the typification of social choice and social preference into a system of hierarchically ordered types can be useful for the understanding of the economic aspects of the process of choice and valuation. But the danger in adopting such model is losing sight that individual preferences are not reflected upon, in experience, in the way which formal models in political economy express such preferences. The adoption of “neutral” points for the distribution of ideal preferences in social plans (and Schutz is particularly critical of Pareto at this point) is further evidence of what we lose when we adopt such methodology⁹.

Schutz provides a different methodology in which we can deal with the constitution of economic and social preferences. I will try to use a game developed by António Damázio and Antoine Bechara¹⁰, the Iowa Gambling Task, as an example of a decision experiment that can be interpreted under the methodology suggested by Schutz:

The Iowa Gambling Task was developed as a tool to present the hypothesis of the somatic marker. This hypothesis, which is the central point in Damázio’s classical study on the brain Descartes’ error¹¹ is that what we call a “decision” is such a complex and detailed process within the brain that it is impossible for an individual to consciously grasp all that is going on when she thinks she is deciding upon a preference or upon a strategy.

Damázio defends that the ventromedial prefrontal cortex (vmPFC) centralizes the input received from the “perceptive areas” of the brain, including the amygdala, the olfactory system and the temporal lobes, and then sends signals, or somatic

⁸ Alfred Schutz, *On phenomenology and Social Relations*. (The University of Chicago Press: Chicago, 1970), 144-5.

⁹ *Ibid*, 133

¹⁰ Bechara A, Damasio H, Tranel D, Damasio AR (1997). "Deciding advantageously before knowing the advantageous strategy". *Science* 275 (5304): 1293–5.

¹¹ António Damázio. *Descartes' error. Emotion, reason, and the human brain*. (New York: Putnam, 1996).

markers, to the “cognitive areas” of the brain which will further interpret these outputs as “emotions” or use these outputs to conceive of “strategies”. However, since the vmPFC centralizes the entire input of perceptions given to an individual, it also works as a “filter” of information for consciousness. What we later recall as the core of relevant perceptions for a determined experience, or for a decision, is already somewhat filtered within the vmPFC.

The Iowa Gambling Task aims at testing this hypothesis. The test asks an actor to choose among 4 different decks of cards A, B, C and D. Each card within the deck is labeled with an economic value. Each deck contains a different harmonic distribution of advantageous or disadvantageous cards, so the harmonic distribution in deck A will be for eventual marginal gain, short term significant loss and long term significant loss (in game theory, this will be a mini-mini situation), B will have a harmonic distribution of eventual significant gain, short term marginal loss and long term marginal loss (a mini-max, in terms of game theory), C will have a harmonic distribution of eventual marginal loss, short term marginal gain and long term marginal gain (a max-mini, in terms of game theory), D will have a distribution of rare marginal loss, short term significant gain and long term significant gain (a max-max, in terms of game theory).

Individuals enter the game in a blind situation: they do not know that the decks have this composition. All they know is that they have to choose a card within each of the decks, and that they have 50 choices to maximize their gain. However, for every choice of a card within a deck they are informed by the simulation that they have lost “x” and won “y”. If the result is positive, the simulation provides the actor with positive feedback (a smiley face on the computer screen and a pleasant sound, if the test is conducted by software, positive reinforcement, if the test is conducted in person). As individuals play the game it is expected that they will learn and understand the patterns of each deck, choosing the more advantageous deck in order to maximize gain and avert risk.

And, in fact, this is the most typical result for the test. Individuals indeed learn which deck is more economically advantageous and then establish that it is a dominant strategy to always choose the deck “D”. However, this is the typical result, not the necessary one. And some variation is perceived according to social standing, education level, compulsive behavior and brain damage. Some variation has also been observed according to the place wherein the test is conducted.

Of course, I cannot problematize the implications of this test in terms of brain physiology. It should suffice to say that the results seem to confirm the existence of the somatic markers, particularly when the results of individuals with extensive damage to the vmPFC are so dramatically different from individuals that do not show any kind of brain damage. It should also be said that individuals that express compulsive behavior often show understanding of the task and of its implications, they often know that the deck “D” offered more gain, but they actually preferred to take risks in decks with less harmonic compensation. It is also interesting to perceive that the results tend *not* to change if the compensation is virtual or actual. In the original game, individuals actually gained and actually lost the amount given in the cards. In other scenarios, where compensation to participants in medical experiments is not allowed, the

compensation and loss was only virtual. In both scenarios, the results tend to remain homogeneous: typically, individuals will a) perceive that there is a more advantageous deck, and b) choose from that deck.

This test is often taken as a confirmation of most rational choice theories: individuals indeed organize choices in terms of gain, individuals indeed prefer to take a dominant strategy which is advantageous rather than a strategy that is not, and individuals often behave oriented towards goals.

I think, however, that we can try to read the results in a different manner.

First of all, we need to understand how the mind is operating while it plays the game. Surely, for the economist the moment of decision is isolated in the moment in which we “play”, or in which one actually takes a card from a deck in one’s hand and sees the result. Of course, this could be isolated as a “moment” of decision, but I think that Schutz can provide us with a more complex reading of this phenomena.

I do not want to argue (and I do not think I could, in good faith) against the hypothesis of the somatic marker. Actually, I believe we can use it to strengthen the point of a phenomenological reading: I have been stressing throughout this section that the main element of a phenomenological reading is the perception of external information which is affecting our processes of constitution of something as a “concept” or the way in which we value one thing rather than another. I also stressed that much is lost when we remember, or when we recall a presentation forth as a memory. We are not fully in control of all the elements that constitute a choice. We cannot rank all these elements; we cannot really understand how they affect our thought process. And yet, they are affective.

Somatic markers seem to be doing precisely this: they seem to be providing consciousness with some elements that will be apparent to us in the moment of a decision, while at the same time a number of other affections continue to inform the constitution of something as a choice. In that sense, the “ranking” of preference or the establishment of a dominant strategy is but an aspect of the entire process leading to the choice. It is a standstill image within a movie.

Schutz, then, is not very far from current hypothesis in decision theory, particularly to what Damásio has been stressing, when he writes that the problem of social action should be understood as a question of perception and temporal constitution of perceived data. Social phenomenology, in that sense, provides us with leading clues that we can use in order to interpret economic action. In the examples of the Iowa Gambling Task: should we not consider the constitution of a “gain” for the player?

If we subscribe to a purely economic theory of choice and decision we will have a very direct answer to this question: individuals that do not aim at economical gain are not acting rationally. There is something fundamentally wrong with their conduct if it is a conscious conduct towards consistent economic loss.

Surely, we would all be surprised if we witnessed an individual burning money on the street. But we make uneconomical choices all the time, we choose against our economic interest often and consciously (anyone who has opted for an academic career has probably made this kind of choice consistently and

repeatedly). This goes to show that the notion of “gain” is not only connected to the apparent economic advantage of a choice (even though they are analogic to burning money on the street). Rather, a choice is connected to a system of previous affections and references, to a stock of presentations which is represented as a choice *consciously* – but the elements that are represented rely deeply on a number of other elements that affect the decision passively, elements that are not ranked and that cannot be expressed as a “reflected upon” decision.

In that sense, I dare say that the max-max in the Iowa Gambling Task is dependent on socialization more than economic gain. Individuals might very well perceive more gain in choosing the deck “B” because they like the color red so much, and the deck B keeps on providing them with cards in the color red. They might have been socialized in a context where getting a red card is good luck (even though they are suffering staggering economic losses by keeping on choosing the deck “B”). Perhaps the negative reinforcement given by the computer in the software simulation (a sad face and a strong beep noise) is pleasurable for that individual, and she hence insists on that sensation (though that sensation might be costing her real money). Our stock of presentations might inform our conception of gain and utility differently, and this should motivate us to look at expression of sets of preferences that ignore the existence of the stock of presentations or that insist that we should abandon such stock with suspicion. These models will necessarily be artificial and super-imposed. They will also rely on a particular stock of presentations that allows for the consideration of a view of nowhere. This is not to say that such exercises are not useful, or that they are not efficient as a pattern to analyze economic utility. As I have stressed in my first chapter, the methodological soundness of these models is hard to be argued against. But they cannot be framed as an irresistible model of analysis for social choice. Social phenomenology allows us to think the principle of utility differently and in context, thus informing economic theory with elements of disorder and unpredictability – and providing a more complete picture of the decision process.

THE RICHNESS OF EXPERIENCE AND THE POORNESS OF THE CONCEPT

Now, if Schutz has a big point of departure from both Weber and Husserl, it would be in his interpretation of social experience and the conceptual reduction of such experience. It seems that the interlocution with American pragmatism, particularly William James and John Dewey, motivates Schutz to become increasingly suspicious of attempts to establish a transcendental theory of intersubjectivity, or even a transcendental account of knowledge¹².

Schutz ultimately accuses transcendental methodologies of a solipsist attitude, where even other individuals are taken as data available for consciousness. The transcendental ego, in Husserl, takes the existence of other

¹² Alfred Schutz, *On phenomenology and Social Relations*. (The University of Chicago Press: Chicago, 1970), 252-3

egos as mere datum that are interpreted within consciousness, it is not in interaction with others in a social context¹³. Moreover, the phenomenological reduction, taken radically as a transcendental approach, might lead to the bracketing of the existence of others, and the integrality of a self would be hence reduced to its solitary life.

However, this is a brute simplification of the transcendental ego in genetic and generative phenomenology. As Husserl advances his argument we perceive that, first, the lifeworld is *primordial*, it constitutes a condition for the integration of sense, and the ontological structure of these world will therefore manifest itself for everyone, it will constitute the primordial stock of presentations which will further inform a pre-predicative level in which all individuals within the lifeworld will disclose sense. Thus, the pre-predicative level of a lifeworld which is retained by consciousness indicates that external reality is persistent, and that it exists not only for me but for others within that world¹⁴. Moreover, it indicates that I am not only in monadic relation to other monads in a static reality: it points that the solitary life of the self is only a methodological clue into the understanding of materiality, instead of a matter of fact about our selfhood. Actually, the transcendental aspect of our selves, what Husserl calls the transcendental Ego, refers to the privileged status of external reality, and the transcendental pull that objects in that reality operate into our consciousness – in that, Husserl suggests that individuals share a determined structure which allows for the signification of these “pulls”.

So, Schutz misunderstands the problem of origin and primordially in Husserl, reading as a sort of solipsism. But it is interesting to note that as he attempts to provide a phenomenology of the social world which focuses on the natural attitude (in order to avoid the solipsism of the transcendental attitude) he ends up surprisingly closer to a generative phenomenology, and this is clear in his critique of the notion of ideal types in Weber.

Schutz criticizes the notion of ideal types in Weber, but not out of a disagreement with the hypothesis that individuals tend to overemphasize certain aspects of experience and turn them into ideals. His main contention is with the notion that ideal types would somehow enrich or purify experience, working as a filter of sensible pollution. For Schutz, *any* sort of typification is necessarily at loss against the experience of what is conceptualized. As we turn experiences into concepts we lose a great deal of what was experienced – we cannot communicate the whole of a feeling or of a value into a typification¹⁵. Moreover, types will often go against social experience and deny the density of feelings and behavior. The Weberian attitude towards concepts and types, and the sort of social reality these types produce will therefore produce a stratification of reality which turns the process of resignification and historical reevaluation of

¹³ Dan Zahavi, *Husserl and Transcendental Intersubjectivity*. (1996 repr. Athens: Ohio University Press, 2001) 24

¹⁴ Anthony Steinbock. *Home and Beyond: Generative Phenomenology after Husserl* (Evanston: Northwestern University Press, 1995), 193

¹⁵ Michael D. Barber. *Equality and Diversity Phenomenological Investigations of Prejudice and Discrimination*. (New York: Humanity Books, 2001), 107

experience into a history of grammatical reconsideration of types. Thus, the hermeneutics provided by Weber are located on the level of the normative comprehension of society alone, they do not deal with the actual uses and appropriations of experiences in everyday life.

When Schutz describes this process he is in fact dealing with the problem of the origin of sense, and the generation of new meaning within different social worlds. This is precisely the point of a generative analysis of the lifeworld, where sense is not only generated within solipsist acquisition of external data, but out of a familiar world which includes heritage, tradition and so on. From the standpoint of a generative phenomenology, then, Husserl and Schutz seem to be in the same position regarding types: the normative level of typification loses the richness of presentational experience¹⁶, it also points at the constant level of transformation and re-consideration of what is a tradition, how it is constituted and re-constituted within different societies.

Typifications then are useful only insofar they operate as a scheme of orientation for members of a determined society. They create a common grammar that allows for a normative organization of society¹⁷. However, these processes of organization take the social organization of a social group itself for granted, as well as the formal structures that constitute such social group.

THE REFERENCE OF SOCIAL ACTION

Weber had to stress the prevalence of meaning in his study of social relations. This is particularly clear once we understand his methodology: social actors operate in society in a diverse fashion, but sociology can only be occupied with the aspects of this operation that can be circumscribed to rational expression of values and preferences by this individual. Weber trusts that society is ultimately formed by rationally oriented individuals that pursue their goals in a conscious and ordered fashion. In that sense, when we act in society we act with the meaning of our act fully constituted, at least from our point of view, and we expect our interlocutor to be able to also understand the meaning of our action. Accidents in this process, as far as Weber is concerned, are either a result of primary aspects of socialization that were not shared, or of instrumental interpretation of social acts for aggressive motivations. Still, the orientation of the action is towards meaning. We “mean to say” something when we act in society, we want to communicate an interest and be understood in this interest.

If Weber focuses on the meaningful constitution of types, Schutz wants to privilege the experiential aspects of these types. For Schutz, we are always inserted within a circle of expectations within social reality, and this circle of expectations already informs the reception of our social actions before we actually express any preference. More tragically, our points of view will never meet the expectations and subjective view point of the other person that we refer

¹⁶ Ibid, 195

¹⁷ Alfred Schutz, *On phenomenology and Social Relations*. (The University of Chicago Press: Chicago, 1970), 121

to. Expressive movements, intonations and erotic interactions in general enter interaction along with what we actually mean to say, and further complicate the picture of an action as oriented towards meaning. Rather, social action presupposes a sort of mutual attraction and a will to leave the solipsism of a first person perspective and to enter social interaction.

The core of Schutz interpretation of social interaction, then, is based on an interpretative effort to disclose the peculiar motivations of individuals (or groups) as they enter social relations. Not only does he focus on elements of meaning (or appropriation of meaning) within these groups, but also on how their common lifeworld has been constituted. In that sense, even under a radical phenomenological reduction, there is no way one can leave the position of a point of view when considering social action.

Interpretative sociology, as coined by Schutz, will then focus on analyzing social interaction and the uses of terms in social interaction as contingent. It also recognizes a certain disorder in the processes of signification and appropriation of tradition and experiences.

RE-CLAIMING A TRANSCENDENTAL APPROACH TO SOCIAL REALITY: THREE SHORT PROPOSALS

Perhaps a phenomenological analysis of meaning constitution could help us understand why certain societies recognize that women have the right to drive, and certain societies do not allow women to drive, but it does not allow us to say that allowing women to drive is fundamentally more rational and legitimate than not allowing them to drive.

Habermas¹⁸ stresses that phenomenology remains on the level of mere description of social reality, if it chooses to follow the terms of Schutz' interpretative sociology, never really analyzing how the social reality that is constituted has a fair or unfair systemic organization. Even worse, if we are to take a more transcendental approach, and choose to focus on the structure of the self and the presence of a field of affections that restrict and constrain meaning, then we de-rationalize the process of meaning constitution, leaving the possibility of a rational constitution of a systemic order aside, and, with it, abandoning the project of modernity.

If Habermas is right, then, phenomenology is not compatible with any modern or liberal account of political dynamics within society, and it hence denies the necessity of the regimen of rights. Consequently, a phenomenological position will deny the language of human rights as anything but a historical construction, one that has the same degree of validity and normativity as dogmatic and totalitarian conceptualizations of social preferences.

I want to further present three general theses that I take to be compatible with both a phenomenological perspective *and* a modern account of political

¹⁸ PONTIN, Fabrício. "The invention of a paradigm: the normative conception of the lifeworld in Habermas". In TRAVEJTELOVÁ, Jana. *The yearbook on the history and interpretation of phenomenology 2015*. pp. 27-65

philosophy. Since I have been stressing the problem of social choice and decision throughout this dissertation, I will attempt to delimitate the implication of my theses to this issues.

A. CULTURE IS A STOCK OF PRESENTATIONS

Interpretative anthropology indicated that culture should be understood in terms of a web of significations that inform and constitute social relations. This interpretative standpoint for culture owes a great deal to Weber, and sticks to a regional understanding of these webs of significations. Geertz defends that we should understand and interpret signs in the network of significations where they are active, in a way, making ourselves more familiar with determined uses of terms and signs that will vary from place to place.

Geertz uses the distinction between an eye twitch and a wink to illustrate his point¹⁹. From the standpoint of mere observation, it is quite difficult to distinguish which is which, and anyone who has been in a situation of attempting to court someone else might have asked the question ‘was that a wink? Or a twitch?’, and faced the disastrous consequences of a misinterpretation of the sign. Part of the difficulty in interpreting a body sign is not only in distinguishing between involuntary movement and voluntary movement, but also in distinguishing on what is meant by a determined movement. When I wink, do I communicate an intention to flirt? Do I communicate that I am lying about what I just said? Do I communicate to a peer that I wish to conspire against someone? Do I wish to threaten an opponent in a card game? The codified meaning of the act of winking is hence dependent on a number of aspects of socialization that integrate that act into a web of references to a cultural stock of knowledge, actually, it is also dependent that the act is rendered *as meaningful*. After all, certain gestures that are charged with meaning in one cultural scenario might be completely meaningless in different contexts.

Geertz brilliant analysis of culture, however, remains on the level of a stock of knowledge to which an interpreter refers to when attempting to clarify the meaning behind certain acts. Surely, interpretative anthropology goes above the limits of structural anthropology, and shakes some of the most static presuppositions behind Levy-Strauss’ analysis of practices. This, however, is not the place to dwell in these distinctions. What I am most interested in is that even from the stand point of interpretative sociology, culture refer to a process of signification which says that we have some control over what the terms we are referring to mean (at least what they mean *to us*) and how we want to control them. This places culture on the level of representation: it already presupposes that we associate things with their meaning, conducts with determined values, and so on.

But I want to question if we are indeed in control of the process of signification and association which is going on in culture: are we actually referring to a stock of *knowledge* when we build and interpret social reality? If

¹⁹ Clifford Geertz. *The Interpretation of Cultures*. (New York: Basic Books, 1973) p. 6-7.

we assume so, we also must admit that the creation of a social reality and the consequential choice of values and preferences leading to the institution of a systemic order are fully under our control: we actually know how and what we are choosing and preferring.

And yet, we are clearly in a conundrum: what does a stock of knowledge refer to? A stock of knowledge may not appear by *fiat* in a subject, it is somehow acquired in social life and in personal experience. Here, Rawls and Habermas will provide us with different solutions to this problem, which I see as a problem of control. For Rawls, such stock of knowledge should be understood in terms of plans of life and comprehensive doctrines that are mitigated within the terms of political liberalism in decent societies. For Habermas, this stock of knowledge hints at the necessity of communicative reason: individuals are prone to organize their social preferences according to meaning-oriented practices of speech which are universally shared by any individual. If societies are not organizing themselves according to the necessity of such procedure, it is because an instrumental approach to reason has taken over the public sphere. For Habermas, communicative reason frames the limits of rational cultural claims, and for Rawls, the only aspect of culture, as something that informs a comprehensive doctrine, which is meaningful for a political perspective is the aspect that will impact the sustainability of a political doctrine aiming at developing fairness, which is to say, cultural claims are only problematic in the moment they impact the possibility of other culture claims to occur at the same time.

Is Weber in a better position here? The historicist approach has a certain malleability that both Rawls and Habermas lack. Rawls lacks such malleability, it seems to me, on purpose: he wants to decrease the scope of analysis of political philosophy to a minimal account of social choice and preferences. Habermas, however, takes the Weberian analysis to its limit, by choosing to turn a regional analysis of meaning and signification into a universal pragmatic that rules all sorts of rational constitution of meaning.

Still, even Weber insisted that we could only comprehend social action in terms of meaning orientation. This is to say that the aspects of culture that matter for our analysis of action and preference should be restricted to meaning-oriented performances by social actors.

But, again, it seems to me that we need to ask how the constitution of such stock of knowledge, which is the reference for meaning-oriented performance, becomes the case to begin with. I want to suggest that the cultural stratum is not only connected to a web of meaning-complexes that are signified and re-signified in everyday practices. Rather, it seems that culture is precisely that which affects our meaning-like assertions about the world, while at the same time not being perceived as affective.

Technology provides us with a good clue to this hypothesis: the use of virtual environments has become “naturalized” among young individuals in most developed countries to the point in which virtual experiences are rendered with a sort of density that defies the arm-chaired distinction between ‘virtual’ and ‘real’ environments. In that sense, virtual environments enter the realm of ‘culture’, as they participate and compose part of the complex to which we refer when we

act. But the key question here is if they are apprehended in terms of knowledge: that is, are we in control of the ways in which technology shapes and orients our conducts? It seems that these elements are affective much before we actually appropriate them as “meaningful”, our interaction with devices, with tools, has always shaped and changed the ways in which we comprehend reality, but I do not think we can claim that we are in full control of the “meaning” these interactions have – and yet, they are part of culture.

Culture, then, refer to the sets of references to which we are exposed, and that constitute our everyday experiences on a passive level. This level is relevant to what we will later represent and apprehend as meaningful. What I want to stress here is that the level of primary socialization, as developed by Weber, indicates that we do not originally refer to a stock of knowledge, that is, an organized web of meaning complexes. We only do so much later, on the level of the representation of retended experiences.

The constitution of something as meaningful, then, already presupposes a cultural stock of presentations, which include all elements that impact upon our choices and preferences, and that we do not control. In time, we will be able to re-constitute some of these elements, and then indeed re-appropriate them in meaningful terms. We can look back at former choices and identify elements in those choices that we were not aware of by the time we established a preference.

This allows us to understand the malleability of the process of choice in more radical terms: our processes of signification and re-signification always points back at a dynamic cultural set that affects us and pushes us towards these processes. The difficulty of a static view of culture, particularly one that views it as related to a stock of knowledge, is that it will ignore that the dynamic incorporation of new elements into the cultural stratum does not allow us to immediately conceive of these elements as meaningful, even though they affect our social interactions decisively.

Consequently, the process of acquisition of knowledge is in a time-lag between “the realization of something (object), of someone (subject), and of meaning (order)”²⁰, and in tension with the reverberation of elements in objectivity, intersubjectivity and order that are not at all under our control, in the shaping of the cultural strata that affects our attempts of constitution. Meaning constitution, whatever the sort of meaning constitution, from the standpoint of a phenomenological account, is hence anachronic in any case²¹. We have a temporal disorder between the moment in which we perceive and are affected by something, and the moment in which we attach meaning to that same something. Surely, we are also able to identify something in a stable form, and then realize the repetition of a certain input. But in this we are repeating our first representation of a given experience, and remain in time-lag.

In that sense, claims of particular privilege of preferences regarding originality or tradition should be regarded with a great deal of suspicion. We do

²⁰ Bernhard Waldenfels, *The Question of the Other*. (Hong Kong: The Chinese University Press, 2007), 55.

²¹ I own this point to Anthony Steinbock.

not have any access to “the way things were” or to the original and primordial stance of meaning constitution, particularly because on a primordial level meaning might not even be possible.

I realize this might get us dangerously close to an overly relativistic account of social preferences, but this is not my intention here. Rather, I want to point at claims of cultural, traditional and moral superiority as fundamentally absurd and indefensible from an epistemological point of view. Surely, we can understand these sorts of preferences as an of appropriation of the time-lag between culture and meaning, one that tries to compensate the disordered nature of the process of meaning constitution by emphasizing some narratives that attempt to provide us with an ordered description of the act of origin. It should be no surprise, then, that the more literal the interpretation of a founding myth, the more dogmatic the values that arise from this interpretation will be.

I would then approach the problem of meaning constitution and the attribution of values in terms of freedom, rather than in terms of culture. Our cultural stock of presentations refers to the basis in which the constitution of meaning is possible; it operates as a dynamic field in which we institute normal and abnormal practices, and this process of constitution and institution of normality and abnormality is further constricted by more or less dogmatic views of culture, which brings me to my second thesis.

B. FREEDOM AS THE POSSIBILITY OF RE-SIGNIFICATION

Consider how your priorities and meaning-complexes would be like had they remained the same ones you had in the moment you were seven years old. On a personal level, I would have to say that most of my priorities would involve having chocolate, going to swimming lessons, and avoiding Sunday school. Interestingly, I now can no longer have chocolate, dislike swimming, and still avoid Sunday school. I suppose I could provide some other examples, to illustrate a certain maturity, but my point is quite simple: we change the way in which we constitute our preferences and how we think about certain objects in time.

Take, for example, the case of a cup of coffee. Now, one can grow up having Robusta coffee of a terrible brand bought in cans at the nearest supermarket chain, and be convinced that one, in fact, hates coffee. In time, however, one might have the opportunity to have a fine, single-origin, Arabica coffee bought directly from the producer, and then realize that, in fact, one loves coffee quite a lot. Of course, the change in perspective does not have to rely on the quality of the bean (though most, if not all, coffee enthusiasts will insist quite verbally on the superiority of Arabica over Robusta); it might rely on the experience of having coffee, as a whole. Allow me to elaborate: it is quite conceivable that one prefers the lower quality coffee because it recalls other sorts of experiences which one associates with that particular coffee. So, one prefers the Robusta coffee, which one is fully aware is of lower quality than the Arabica sort, because that was the sort of coffee one was having during an important event. In that sense, the coffee is not only the quality of the bean, but an entire context, an entire experience which is recalled and presentified in that moment. In both

cases, we have the constitution and re-constitution of optimal appearances of something to consciousness, though in the second example the earlier experience remains optimal even though the individual experiences another “norm”.

But my point here is not about coffee, it is about freedom. How so? I want to hold that the central point for a phenomenological contribution for political philosophy will have to retain a privileged space for a notion of freedom as freedom of re-signification. Of course, as long as we are talking about coffee, we are relatively free of dogmatic views about changing one’s mind. But what about Sunday school?

Chocolate, swimming and coffee do not offer us much controversy in terms of practices. Few people actually mind that one does not like coffee, swimming or chocolate, and even fewer people seem to be willing to kill you if you change your mind regarding a preference for Arabica to Robusta. Now, if we look at more controversial values, the situation changes immensely.

Reconstitution of preferences and values regarding organized religion, ethnicity, sexuality, and other aspects that are perceived as absolutes within different cultures seem to be problematic. And yet, they are a central part in the development of societies that tend to be less brutal than others.

In that sense, the impossibility of re-signification of values, grounded on the supposed originality or superiority of a certain claim of culture, institutes a very peculiar kind of bondage that inflicts a direct impediment of speech.

Freedom of speech as freedom to re-signify and re-appropriate becomes central to my phenomenological interpretation of the political: the process of meaning constitution is always on the open, and opens a venue in which we can articulate and re-appropriate culture positively or negatively, we can access former experiences, or current cultural norms, and deem them racist, inappropriate, undesirable, and so on.

Now, it would be naïve to suggest that such process will always be successful. It will not. It is quite possible that in time conducts will be re-signified and an individual might become *more* racist or *more* totalitarian. This is, in fact, quite a common response when we look at totalitarian social groups that feel threatened by the appearance of “anomic” competing individuals or groups.

This also points at a comprehension of cosmopolitan history outside the limits of a teleological and evolutionary base. From a phenomenological standpoint, peaks of mutual recognition and understanding might be followed by disastrous new norms of conduct. We are not inexorably moving towards a more cosmopolitan political order.

Interestingly, this takes me back to some of the conclusions of the analysis of democracy in Arrow: if we look at how democracies fail, part of the problem is with whom feels motivated to organize and speak about values. As a democratic order becomes more organized and stable, individuals tend to take the institutional framework for granted. In that sense, a stable and relatively apathetic relationship with the process of voting and active political participation might be (and should be) interpreted as a good sign. General commotion towards elections usually indicates incipient or failing regimens (think of the 2011 protests at the Middle East, or the commotion of the 1989 presidential elections

in Brazil – the first direct election since 1960, and the first fully inclusive election in the history of the country). But the paradox in this situation is that the individuals that will feel motivated to speak in these elections, that will actually vote, are individuals that do not represent the majority of society. These individuals benefit from an order that allows them to claim for particular policies, for particular re-considerations of systemic norms, and then might use these norms to *decrease* the reach of freedom of speech and re-signification.

This puts us in an awkward position: attesting the freedom to re-signify values and preferences might cause us to, in time, *decrease* the possibilities of signification: it might move us to a totalitarian position of speech where a determined norm, a determined view point about society, dominates all forms of assertions.

We are then again faced with the problem of control: how do we control assertions about preferences in order to, first, recognize that we cannot simply establish an ought-like structure for social relations, and, second, that an uncontrolled system of assertions about preferences tends to anarchy and totalitarianism (often at the same time).

My final thesis about this matter is quite liberal: I want to defend that we need to think politics through as a system of organization for our processes of signification.

C. POLITICAL ORDER AS LIMITING THE SCOPE OF SIGNIFICATIONS

Rawls' insight about the toleration of the intolerant is one of the most compelling examples of how to defend freedom of speech while retaining some sort of coercion against violent behavior. For Rawls, the key point is the right to complain: the right of complain is a fundamental principle of justice, which follows from equal liberty. Until an intolerant sect gives us reason to perceive it as threatening our basic civil institutions, or our physical integrity, we have no justification to “curb” the intolerant²². Unreasonable claims of any sort are acceptable insofar they are understood as a manifestation of a shared equal liberty among all individuals: the liberty to speak one's mind.

“Threat”, is, nevertheless, a complicated concept. I am sure that Rawls understands it in terms of direct and expressed intent to destroy civil liberties, and direct and expressed actions towards violence against individuals. But what about symbolic violence and symbolic threat? Are those levels irrelevant?

I do not think that Rawls disputes that language plays a role in constituting social reality and the fairness of institutions, but, and this was an important claim in my first chapter, it seems that his focus on political economy as providing a neutral point of view for the development of the original position suggests that interpretations about the symbolic violence of a determined social actor are only relevant as they are in conflict with normative guidelines.

²² John Rawls, *A Theory of Justice*. (1971; repr., Cambridge:Harvard University Press, 1999), 190-4.

Granted the scope of Rawls' analysis, I do not have any reservations against it. Indeed, from the standpoint of a normatively constituted order, we have much more to lose than to gain by regulating speech excessively. But the problem here is precisely with the scope of the analysis: how do we get to the point where we have a normatively constituted order?

If we follow the three thesis I am presenting progressively, we have (1) a cultural stock of presentations which is (2) consistently appropriated and re-appropriated meaningfully in individual and social practices and experiences that (3) will need to be regulated, in time, in order to avoid a complete collapse in social relations. In that sense, I hope that a systemic level will serve as a control for the processes of signification and re-signification of culture, thus dramatically increasing the scope of the analysis provided by Rawls.

The problem with increasing the scope of analysis here is avoiding falling prey to the same issue that plagued Habermas: the insertion of an ought-like structure for social relations. First of all, I think I already dodge this issue by introducing a non-teleological account of the process of constitution: I admit that the process of constitution and re-constitution of preferences is in the open as it does not follow the necessity of communicative reason, it only follows clues given by a cultural stock of presentations.

But even if you grant me that I have avoided the problem of an ought-like description of social action, I still need to clarify how norms are possible within this perspective. I think that I have already anticipated much of these movements in the introduction of this article, when I brought up the question of optimality and normality in Husserl. But I did not offer any description of these processes on an institutional level in Husserl.

This is because Husserl offers nothing of that sort. We will not find in Husserl any direct reference to the constitution of norms on an institutional level, and how these norms influence and change social behavior. However, I do not think it is impossible to incorporate these elements onto the "stock of preferences" that influence the way in which we make decisions.

Allow me to suggest that we can interpret a movement within Weber phenomenologically in order to have a clue into this issue: grant that we have a disordered set of references which is informing the constitution of normal and abnormal conducts. Now, once this disordered set is interpreted within a familiar context, we start to conceive of what Husserl calls a "homeworld", and the formerly disordered set of references is organized within a system. At this point, the cultural stock of presentations becomes a stock of knowledge. In time, our social interactions will move us away from our initial topological position, disturbing what was previously organized as "normal" and "abnormal", this generates new possibilities for conceiving of preferences (as in my example with coffee, or, in a more radical example, religious conversion), and also increases the scope of the stock of *presentations* to which any future meaning-like conceptualization will ultimately refer to. Now, this is the point in which conflict between different sorts of conception of "normal" conduct will become more radical, as they are going to refer to fundamentally differently constituted stocks of knowledge, that is, each "home" will constitute the conflicting form of normal

constitution as “alien”. Once again, the social interaction here seems to point at the necessity of implementing a grammar for social relations which will increase the scope of familiarity, but the formation of such grammar is going to be effective within a context where the reference point of discussion has been affected by this irreversible encounter with an alien perspective. In that sense, the formation of a more or less totalitarian State which will regulate future mutual encounters is directly related to the history of previous interactions with different perspectives. A cosmopolitan position, in this context, is only possible within a perspective that will give the historical conditions to such possibility.

Then, if we want to claim that the historical movement leading us to the current discourse on human rights is indeed irresistible, we will have to find ways to inform and communicate the necessity of such approach within communities that may have reason to be suspicious of our suggestions. This is not to say that we should drop our historically constituted preferences about the good, but to put them in perspective, including how the pathway leading to the prevalence of a discourse on human rights in developed democracies has often and consistently been implemented on the basis of the exploration of underdeveloped countries. The history of the formation of the Westphalian consensus, in this context, is particularly informative, as the development of the way of life in the European Empires, leading to the implementation of a discourse on rights, was only possible while the same European potencies that claimed the values of human dignity were, and still are, exploiting western Africa to the point of destruction.

In this sense, the process leading to the legitimacy of modern democracies and the universal claim of Human Rights, which I also believe is a claim that finds full support within the context of the Husserlian later philosophy and its obsession with the relevance of European humanism, can only be implemented on a broader level in terms of an exchange of historicities, of a mutual awareness of the processes leading to current mainstream conceptions of goods in different States and communities.

VALUE AND MORAL PLURALISM

FLÁVIA C. CHAGAS

“Some people cross the forest and see only firewood.”
(Leo Tolstoy)

The idea that value judgments are subjective, lacking objective theoretical warranties, or impossible to be understood in a rational justification is not new, not only in philosophy, but also in common sense. Actually, not touching upon sociological, anthropological or psychological—in other words, empirical—discussions concerning whether and how it is possible to morally justify our choices having in mind possible factors that condition (but not determine) our actions, philosophical skepticism generally supports that it is not possible to dispute valuation questions so that questions about “what is good?” or “how should we live?” concern only the subjective dimension of moral evaluation and decision.

As is well-known, the Kantian perspective goes against the kind of argumentation outlined above, in which we may include both skeptical perspectives, as that of the amoralist, for instance, and emotivist perspectives, which reject the possibility and even the necessity of justifying moral values and norms.

However, the thesis that practical reason should act as a mainstay for the justification of moral values is not restricted to the body of Kantian universalist ethics, but is actually defined among wider theoretical perspectives, so to speak, as the case of Hilary Putnam’s pragmatism or of the social liberalism endorsed by the philosopher and economist Amartya Sen. Indeed, in the beginning of his book *The Collapse of the Fact/ Value Dichotomy and Other Essays*, Putnam posits that

The idea that “value judgments are subjective” is a piece of philosophy that has gradually come to be accepted by many people as if it were common sense. In the hands of sophisticated thinkers this idea can be and has been developed in different ways. The ones I shall be concerned with hold that “statements of fact” are capable of being “objectively true” and capable, as well, of being “objectively warranted,” while value judgments, according to these thinkers, are incapable of object truth and objective warrant. Value judgments, according to the most extreme proponents of a sharp “fact/value” dichotomy, are completely outside the sphere of reason. This book tries to show that from the beginning

these views rested on untenable arguments and on over-inflated dichotomies. And these untenable arguments had, as we shall see, important “real world” consequences in the twentieth century (Putnam, 2008, p.13).

Indeed, according to Putnam, the idea that value judgments are strictly subjective and, therefore, cannot be rationally disputed results from certain philosophical stories and constructions that began with modern empiricism, but were reaffirmed and consolidated with their 20th century “offspring,” namely logical positivism. Thus, one of the objectives of Putnam’s analysis in this book is to show that the reason why value judgments lack objectivity is found in the advocates for the dichotomy between fact and value, which is, on its turn, attributed to the well-known Humean argument that from judgments about how the world is (descriptive enunciates about facts) it is not possible to infer how the world should be (prescriptive enunciates about values). But this distinction echoes another, namely the distinction between judgments about relations between ideas, known as analytic judgments, and judgments about questions concerning fact and existence, known as synthetic (*a posteriori*) judgments, so that such judgments synthesize the set of enunciates capable of carrying some kind of objectivity, since the truth value of these enunciates depends upon either the simple logical relations among the concepts involved or the possibilities of assessing the correspondence between representation and object.¹

¹ Putnam does not reject the necessity and the utility of the distinction between fact and value. What many theoreticians, among them Putnam, Sen, Nagel, Hare, Habermas and Korsgaard, reject is the thesis that, to be objective or to have some kind of pretention of objective validity, judgments should be descriptive. Indeed, Cristine Korsgaard points in that direction when she posits, in her article “Realism and Constructivism in Twentieth Century Moral Philosophy,” that “what is important for my purposes is this: even when what we might call the verificationist element in verificationism was dropped—that is, even when philosophers reclaimed the intelligibility of propositions that cannot be verified through the empirical sciences—one element in the verificationist picture was retained. That element is the idea that it is the function of all of our concepts, or anyway all of our authentically cognitive concepts, to describe reality. We must go carefully here. In calling this into question, I don’t mean to deny that there is a sense in which all of our concepts—that is, all of the concepts we have any business using—can be used in propositions that do in fact describe reality, in the sense that they are capable of being true or false. Rather, I mean to call into question the idea that this is what all of our concepts are *for*—that their cognitive job, so to speak, is to describe reality. So long as we retain that idea, it will continue to appear that moral realism is the only possible alternative to relativism, skepticism, subjectivism, and all of the various ways that ethics might seem hopeless. And so long as moral realism appears to be the only alternative to these skeptical options, the need to show that moral truth is as solid, as real, as objective, as scientific truth—and also that it is objective *in the same way* as scientific truth—will seem pressing. This was our situation in the early and middle years of the twentieth century. I am aware that what I have said about the function of our concepts will seem vague until I articulate an alternative function. I intend to do that presently, but first I want to notice that from the start there was already a problem with the distinction between cognitivism and non-cognitivism in ethics. The distinction suggests that a moral judgment either articulates a description of some fact or is as a disguised version of some alternative use of language—either expressive or prescriptive. But where does this leave theories like Aristotle’s or Kant’s, according to

Without entering into the details of the difficulties among the truths by convention and Quine's critique of the analytic/synthetic dichotomy or in the problems of the defined descriptions and the question of the notion of truth understood as correspondence, it is important to define the limits of our reconstruction of the problem of the justification of ethics so as to draw attention solely, at this moment of the argumentation, to the fact that not only tautologies or factual/empirically verifiable truths may reclaim objective validity for its enunciates, but also certain enunciates about moral value without the need to address the problematic notion of moral facts or truths, which would have to be, so to speak, discovered.

Does saying that philosophical ethics does not have the function of describing moral facts or truths that could, in this case, be the object of a description, imply the impossibility of disputing value questions? Or yet, is the renunciation of objectivity in ethics and in questions of value based on the fact that ethics does not have a descriptive function? And, moreover, is the solution inevitably skepticism in terms of values?

If we understand the task of ethics as moral philosophy not only, or better, substantially, as that of describing the world, but as that of intending to offer a possible guide for human conduct, that is, an answer to the question about the good life, it seems impossible to deny the possibility, and even the necessity, according to Kant and Putnam, of disputing not only the means, but also the ends when we treat valuation questions. However, this dispute, according to these philosophers, depends on and goes through finding a notion of rationality applicable to normative questions.²

This notion of rationality should be capable of guaranteeing the moral pluralism of several conceptions of good without, on the one hand, falling into

which moral judgments are the conclusions of practical reasoning? A conclusion of practical reasoning is not obviously a description of a fact about the world, but it hardly seems like some sort of emotional expletive, either. Where do these theories fit? I believe that the answer is that they *don't* fit, but unfortunately this may not be obvious, because the issue raises a further question about the principles of practical reason themselves. We may say that it is true that an action is right just in case it accords with the categorical imperative, for instance, but what then are we to say about the categorical imperative itself? When we ask about the status of the principles of practical reason, the question of cognitivism and non-cognitivism seems to come up once more. For instance one may suppose that the principles of practical reason must be self-evident truths known by intuition, and then Kant will come out looking like a traditional rationalistic realist. This is not just a fantasy—this is how Kant was actually read by many late nineteenth-century philosophers, especially in Britain. Sidgwick and Mill, who read Kant this way, will serve as sufficiently distinguished examples. Alternatively, someone who takes seriously Kant's thesis that moral laws are the laws of autonomy, legislated by the agent's own will, may read him, as Hare sometimes seems to do, as a prescriptivist and so a non-cognitivist. So although practical reason theories might at first seem to fall between the cracks, there are ways of making them fit the mold." In: *A Filosofia Prática de Kant: ensaios*. Pelotas: Nephil Online, 2014, p.25-26

² This suggests, beforehand, the skepticism in relation to the set of theories that deny that practical reason plays a fundamental role in moral justification, such as current expressivist and non-cognitive theories

moral relativism and, on the other hand, searching for an ultimate groundwork for what is morally good as understood in terms of a “strong” moral realism, such as is the case, for instance, in Boyd and Railton, who advocate the general idea that objectivity in ethics supposes the notion of objectivity analogous to that of the natural sciences (that is, that there are moral facts that are objectively valid or independent from the mind).

Thus, the aim of this paper is to show that, and how, the Kantian perspective is capable of offering, if not the best, at least a good option for solving the problem in question, namely finding and justifying the notion of practical rationality capable of warranting moral pluralism without falling into relativism.

* * *

Although the objections to the Kantian paradigm are well-known, I would like to draw attention to some that can be productive to the dialog in this text: 1) that it is a timeless, de-contextualized and ahistorical notion of rationality³; 2) Kantian ethics is an ethics of the “duty for duty’s sake,” which means a rigorist ethics that demands too much from the subject⁴; 3) it is an ethics that does not include the pluralism of conceptions of good,⁵ since the only ethics capable of having moral value is that founded on the CI.

I will try to show that these three objections are unjustified due to the reasons that will be developed throughout my argumentation: 1) the claim that it is a de-contextualized and ahistorical notion of rationality depends on assuming a theory of the two worlds, or yet, the support of a double ontology to justify this conception of practical reason; 2) as Kant, Sen and Putnam posit, it would be a very shallow presupposition about human beings to identify rational choice with the search for realization only of self-interest, having in mind that the subject values and has reasons to value “things” such as commitments, ideas, values, that are not exhausted in the pursuit of pleasure and the tendency to avoid pain. This point may also call to mind the important link in Kant between the notions of morality and virtue; 3) finally, the third objection is that it is an incomplete ethics concerning the justification of good in ways that can integrate different conceptions of good. The Kantian proposition is exactly the contrary: that only through a formal criterion of evaluation it is possible to ensure moral pluralism without falling into moral relativism.

In order to avoid misunderstandings, it should be emphasized that this paper aims not so much to defend Kant by attempting to refute these objections than to provide a somewhat detailed investigation of an implication, seemingly fairly

³ Cf. for instance, SCHOPENHAUER, A. *Sobre o fundamento da Moral*. São Paulo: Martins Fontes, 1995; WILLIAMS, B. *Moral: Uma introdução à ética*. São Paulo: Martins Fontes, 2005; SEN, A. *A ideia de Justiça*. São Paulo: Companhia das Letras, 2011; TUGENDHAT, E. *Lições sobre Ética*. Petrópolis: Vozes, 2010; KERSTING, W. *Liberdade e Liberalismo*. Porto Alegre: Edipucrs, 2005.

⁴ *Idem*.

⁵ KEKES, John. *The morality of pluralism*. Princeton University Press, 1993. See also the analysis by AUDI, R. *Moral Knowledge*. Oxford: Oxford University Press, 1997, p. 40-41.

obvious, of the notion of practical reason as found in the Kantian paradigm that apparently has not been discussed in depth by the commentators, namely that exactly against such objections, Kant advocates the pluralism of conceptions of good so that this thesis can only be ensured from a conception of practical rationality as the groundwork for the moral value of actions. In this perspective we read in the Anthropology:

Finally, the *moral egoist* is a man who limits all ends to himself, sees no use in anything except what is useful to him and, as a eudaemonist, locates the supreme determining ground of his will merely in utility and his own happiness, not in thought of duty. For, since every other man also forms his own different concept of what he considers happiness, it is precisely egoism that results in [the eudaemonist's] having no touchstone of the genuine concept of duty, which absolutely must be a universally valid principle.—So all eudaemonists are practical egoists. The opposite of egoism can be only *pluralism*, that is, the attitude of not being occupied with oneself as the whole world, but regarding and conducting oneself as a citizen of the world (Kant, Anthropologie, VI, BA 8).

In order to understand how this connection between the notion of practical rationality as the nucleus of moral value is possible and how this conception allows Kant to say that the criterion of the categorical imperative is what allows moral pluralism, we will firstly investigate the demands and presuppositions of the Kantian understanding of practical agency to see whether and how it is possible to say that it allows moral pluralism.

1. PRACTICAL RATIONALITY AND MORAL VALUE

To understand the Kantian thesis that the moral value of an action depends on its being based on the normative demand of practical rationality, two considerations are indispensable:

- 1) practical rationality, in its formal aspect, is the source of moral value;
- 2) to conceive this demand, which unites aspects of normative or substantive ethics, we do not need to suppose super-inflated ontologies, or, in Putnam's words, metaphysical stories, even from the Kantian text.

In relation to the first point, Kant stresses what he understands as the form of the law as groundwork for the determination of the will in the first two sections of the GMS and the first six chapters of KpV from the distinction between acting according to maxims that may be considered valid only subjectively or objectively. This mode of judgment, according to him, is possible only by the disposition of the subject to act based on the representation of rules.

In the two texts mentioned above, namely GMS⁶ and KpV, we read that the categorical imperative consists, then, not in a concrete norm or in a set of

⁶ KANT, I. Werke in Sechs Bänden. Herausgegeben von Wilhelm Weischedel. Wiesbaden: Insel Verlag, 2011. From now on, I will use the usual abbreviations to cite Kant's works: Critique of Practical Reason (KpV), Critique of Pure Reason (KpV), Groundwork of the Metaphysics of Morals (GMS), Metaphysics of Morals (MS).

concrete norms, but in a formal criterion for evaluating maxims. Then, it is not decided *a priori* which maxims are moral or not so that the subject has to evaluate, from his or her own moral consciousness, which maxims could be thought of as valid from a universal perspective.

Even though Kant evidences that the categorical imperative is not a set of empirical norms or even a moral truth, it is important to stress the fact that this is a very controversial point in the exegesis of the Kantian text, since, whereas Kant is often accused of advocating a cosmopolitan universalism or a realistic metaphysics in the sense of advocating the existence of moral facts that could be grasped by intellectual intuition, we should remind the unwary reader about two central theses of the critical-transcendental project, which are: 1) the **limit** of every possible knowledge—we can only know that which can be the object of a spatial-temporal intuition; and 2) the categorical imperative consists in a formal principle of evaluation of maxims or subjective principles.

For short, we can say that, if the rules are represented as only subjectively good, that is, considered valid only for “my” will, then the groundwork for determining freewill consists in the expectation of the sensations of pleasure resulting from the realization of some desired object. Therefore, such maxims may not aim at universal validity, that is, be considered morally good. However, having in mind that good is not to be confused neither with useful nor with agreeable, maxims have the pretension of being considered moral if they can be considered valid for all rational beings in general based on the criterion of the universalizability of the subjective principles of the will. Thence, two questions arise: 1) how can I know if the maxim that “I” consider good can be evaluated as valid for “us” or for “all”?; 2) how is it possible that a being like man may want to act according to these moral maxims giving up the pretension of satisfying their inclinations and particular interests? This discussion reminds us of the theme of moral pluralism. Before that, I would like to provide a brief remark about the second issue above.

The question about who is the moral subject in Kant, or better, if there is an implicit double ontology in the justification of his conception of practical reason is one of the most controversial and difficult questions “within” and without Kantian philosophy so that a promising path is assuming the thesis of the KrV that all that “exists” may be the object of possible knowledge. Then, Kant does not need to support the reality of a noumenic, supersensible I to ground the validity of the moral principle, but actually to suppose as starting point for this task the pre-philosophical “moral” (practical) consciousness of the empirical subject insofar as he is capable of acting from the representation of rules.⁷ That

⁷ Cf. the following passages that corroborate this interpretation: *GMS*, 412, *KpV*, A 14 (n.), *KpV*, A 19, etc. In addition to these passages, *KpV*(A 53) §6 may provide an interesting clue for interpretation, in which Kant affirms, in an analogy with theoretical philosophy, how we can understand the necessity with which moral consciousness is imposed in the moment in which the subject evaluates his or her maxims. Indeed, the fundamental premise of Kantian ethics is that moral consciousness is *a priori* or necessarily imposed and constitutes a way of self-understanding that is not to be confused with any intellectual intuition, on the one hand, and also, on the other hand, cannot be reduced to the mere

is, Kant does not assume a double ontology with the distinction between the two perspectives in which man can represent himself as rational and sensible; it is, as Kant himself posits, an epistemological distinction between different points of view in which the same subject can consider himself.

Supposing that it is a “real” distinction means doing what Ryle calls category mistake. Indeed, in “The concept of mind,”⁸ Ryle posits that the myth of the two worlds was actually more harmful than beneficial to the understanding of the distinction between mental and physical.

Thus, if the supposition that the starting point of Kant’s argument consists in assuming that the justification of the validity of the consciousness of moral law as a categorical imperative is based on the empirical consciousness that an agent has in the moment in which he judges the content of his maxims, of an *a priori* demand or constraint of practical reason, then we do not need to commit ourselves with supporting inflated metaphysical theses.

Finally, it should be remembered that for Kant it is not possible to formulate moral maxims independently of the human mind. That is, moral consciousness, or the consciousness of what I should do is only present when the subject evaluates his maxims; in other words, it arises only from reflection in the act of judging one’s maxims. That is why a child who is not yet capable of judging based on concepts would also be incapable of having the consciousness of what to do. To conclude the Kantian answer, realism and idealism are intertwined in the critical-transcendental system. Then, reality depends on human perception (the way reality is and should be depends on the human perspective—cognitive apparatus).

In his book on the “dichotomy between fact/value,” Putnam follows the Kantian argumentation that both our descriptions and our valuations depend on the way we “build” our values and beliefs from the cognitive structure of the human mind, that is, such ways of “seeing” (describing/valuing) already suppose *a priori* a perspective, a point of view—in a word, some place.

Returning to the first point, if a moral action is that which satisfies the demand of the criterion of universalizability of the maxims of the will, which actions can be considered good? Or yet, how can I know what maxims are universalizable?

empirical consciousness of moral duties. Thus, the consciousness that the subject has of what he must do in a certain situation presupposes not only the practical self-consciousness of his own freedom (*Willkür*), but also the *a priori* knowledge of his self-reflexive practical ability (*Autonomie*). For short, the agent that recognizes that he should do X or that something in a certain situation is morally necessary immediately and originally approves of the validity of morality; then the question “How can I know that this consciousness is real or not?” seems senseless at this moment; or yet the perception that beyond the consciousness of the need of doing X I still need some kind of intuition or theoretical warranty that this representation is not a figment. This means, then, that, before thematizing the content of the moral principle, understood as the categorical imperative, we have to presuppose that every rational human being originally, or *a priori*, recognizes moral duties, which are not to be confused with juridical norms, rules and social patterns of a certain society or culture.

⁸ RYLE, G. *The Concept of Mind*. Hertford College, University of Oxford, 2009.

It should be reminded that there is no *a priori* way of defining what can be said as a moral maxim. Kant does not intend to determine, from the criterion of the universalizability of the maxims of the will, a set of moral “truths” or “norms,” which could be consensually conceived as good to guide actions. Indeed, Kantian ethics defines **what exactly** should be considered as morally good. Thus, insofar as the categorical imperative does not establish a set of rules, but a criterion for evaluating rules, the agent should evaluate based on the formal criterion of practical rationality what he or she should do; therefore, “the right thing to be done” implies the reciprocal recognition of the other as worthy of respect in his or her autonomy and freedom.

Thus, if there is no predetermination in relation to what maxims are moral or what rules should be followed, if each subject should judge from his own consciousness what should be done, would we not fall into a relativistic position, something that Kant would definitely not be inclined to accept in his project of ethical justification?

2. SOCIAL PLURALISM AND MORAL PLURALISM

To avoid misunderstandings, I would like to tell what Barbara Herman calls social pluralism in the second chapter of her book *Moral Literacy*⁹ as “the presence of a society of distinct traditions and ways of life,”¹⁰ from the moral or normative pluralism that seems to be connected to a conceptual (or metaconceptual) demand that aims to justify a concept of what is morally good without supposing the existence of moral facts or truths.

Barbara Herman points out to the fact that in many moral perspectives pluralism is seen as a potential factor not to solve problems of moral disagreement, which would then require the possibility of justification of an objectively valid value or norm. Maybe this is exactly one of the reasons why Kant is seen as opposing pluralism, namely due to grounding morality on one “only” principle. However, what I try to advocate, as Herman does, is that “one of the reasons for employing a Kantian model of moral judgment is that it can acknowledge the distinct claim of local values without regarding them as fixed” (Herman, 2007, p.33).

What seems problematic, however, in Herman’s interpretation is that such discussion about pluralism seems not to be solved with what she understands as social pluralism. My conjecture is that, in this systematic point of his argumentation, Kant seems to reclaim something “stronger,” so to speak, namely moral or normative pluralism, which implies a question of second-order justification.

Although Barbara Herman treats the “practical principles that structure the deliberative field” (Herman, 2007, p. 42), the discussion at this moment of her argumentation seems to be focused more on how such principles can be experienced in our moral sensibility as well as in the demand of articulation with the socialization

⁹ HERMAN, B. *Moral Literacy*. Harvard University Press, 2007.

¹⁰ HERMAN, B; op. cit, p. 29, my translation. This conception of social pluralism in Herman does not seem to be distant from what Rainer Forst determines as the ethical context of normative justification in his book “Contexts of Justice.”

and education of individuals than on an investigation of the justification of the scope of the practical reason.

However, Herman intends to show, from the investigation of rational agency in Kant, how the Kantian model, contrarily to the objection of skeptical authors, may be the basis for a pluralist conception of different ways of life in one same moral community. Thus, if social pluralism concerns the different *ethos* in one same community (or different *ethos* in different communities) which we can approach with a merely descriptive point of view, moral or normative pluralism is linked to the attempt to justify a conception of moral value able to refute moral relativism.

In other words, it seems to us that the problem in “remaining” strictly on the level of discussion of social pluralism is that we do not have a way to justify the reason why some conceptions of good in some moral communities are problematic, which then demands a normative principle that is able to show that and why certain conceptions of good **cannot** be justifiable.

If we go back to the original sense of the principle of universalization, we see that it should not be understood as a moral truth or the true content of ethics, but actually as an indicator of a negative criterion, or yet a limit-concept that serves to “assess” the possibility of coexistence of different conceptions of good in a democratic society of rights. In this perspective, it is not given beforehand which should be the ends to be sought or which are the conceptions of good and values that have a moral content, thus warranting social pluralism. On the other hand, any possible conception of good should be compatible with other worldviews from the standpoint of its possibility of rational justification.

From this interpretation of the Kantian moral criterion, it is possible to justify why certain values of a community, as the practice of corrective rape or female mutilation or also, as Herman points out, sadistic desires such as sexually abusing a person, could not be justified as “good” from a moral point of view, exactly because either they harm certain rights or they do not respect the person in his or her autonomy or are founded on an absurd egoism. In other words, they could not be rationally justified in face of the community of human beings to use Rainer Forst’s words in the context of normative justification of morality in his book “Contexts of Justice.”¹¹

But here we find another problem, which constitutes one of the objections cited in the beginning of the text, namely the Kantian justification of the thesis that moral value is based on practical reason. Why should I act according to a normative demand of practical rationality understood from the criterion of universalization? And, in addition, is this demand not beyond the interests and the motivational set of what human beings are able to have? Although this problem seems to be the source of debates and different interpretations, it seems to us that this is not merely a problem of motivation and moral psychology, but of justification and moral philosophy.

¹¹ FORST, R. *Kontexte der Gerechtigkeit, Politische Philosophie jenseits von Liberalismus und Kommunitarismus*. Frankfurt am Main, Suhrkamp, 1994. Contextos da justiça. Filosofia política para além de liberalismo e comunitarismo. Tradução de Denilson Luis Werle. São Paulo: Boitempo, 2010.

To conclude, I would like to draw attention not to the practical self-contradiction of the subject who does not want to engage in a kind of action that involves certain capacities such as the disposition to act based on reasons and justify them also from reasons, but to some insights found in the Kantian corpus and also in different philosophical traditions:

1) a subject who does not ask himself “how should he spend the time of his existence” (Habermas, Sartre) is a subject that seems in some sense alienated;

2) just as Sen, Nagel and Putnam point out, human beings do not define their lives by the ongoing search for pleasure and by doing only what interests them; but they actually value different things that they have reasons to value, such as character, values, commitments, causes and ideals, affective relationships and bonds etc.

3) we can always ask, as Forst, Korsgaard and Putnam suggest, whether that which is seen as something valuable is actually valuable. For such, we do not need metaphysical stories, but the critical ability we all understand from our practices as subjects inscribed in certain traditions, situated in different normative contexts.

Finally, however, we can say that finding out the extent to which such questions about whether so called normative demand are able to serve as a mainstay for our lives is no longer a typically philosophical task, but something that concerns each person’s own intimacy, and here we see the limits of all moral investigation.

IMPLICATIONS AND POSSIBILITIES OF NATURALISM IN NIETZSCHE'S PHILOSOPHY

FELIPE SZYSZKA KARASEK

The relationship between Nietzsche's thought and philosophical naturalism is currently one of the most debated themes within the studies of Nietzschean philosophy, focusing on the implications, the contributions and the limitations that this relationship may bring for the interpretation of themes developed by the German philosopher.¹ In these studies, the main problem is not the association of Nietzsche's philosophy with philosophical naturalism, but the attempt to contemplate in a naturalistic approach Nietzsche's analyses related to nature, culture, morals and other themes, which seem, most of the times, distant from naturalistic approaches. Thus, the different meanings of philosophical naturalism should be investigated and whether this association is a hermeneutic necessity that contributes to the understanding of Nietzsche's philosophy or it reveals a necessity of contemporary philosophy.² Therefore, this paper analyzes

¹ SCHACHT, Richard. *Nietzsche*. London; Boston: Routledge, 1983; LEITER, Brian. *Nietzsche on Morality*. London: Routledge, 2002; JANAWAY, Christopher. *Beyond Selflessness. Reading Nietzsche's Genealogy*. Oxford: Oxford University Press, 2007; CLARK, Maudemarie; DUDRICK, David. *The soul of Nietzsches 'Beyond Good and Evil'*. Cambridge: Cambridge University Press, 2012; SCHACHT, Richard. "Nietzsche's naturalism and normativity", in: JANAWAY, Christopher. (ed.); ROBERTSON, Simon. (ed.). *Nietzsche naturalism and normativity*. New York: Oxford University Press, 2012; among others. The publication *Cadernos Nietzsche* dedicated its 29th edition to the discussion about naturalism in Nietzsche; moreover, articles about this theme can be found in the journals *Estudos Nietzsche* (PUCPR) and *Revista Trágica* (PPGF-UFRJ). Similarly, there are isolated publications in other Brazilian academic journals focused on this theme as well as books concerning some kind of relationship between Nietzsche's philosophy and naturalism.

² For Rogério Lopes (2011, p. 311-312), "this naturalistic turn in contemporary philosophy, visible especially in English language academic circles, could be indicated as the most immediate cause of the marked interest of Nietzsche's interpreters for this theme, as we can see in articles published in this issue of *Cadernos Nietzsche*. Indicating this immediate cause could, on its turn, strengthen the suspicion that these attempts to associate Nietzsche and naturalization programs that are very strong in contemporary debate could have been motivated less by the questions that were indeed crucial to the German philosopher than by an imposed agenda. This agenda, on its turn, could have sprung from an ever growing pressure for explanation that, on its limit, threatens the autonomy of philosophy, in suggesting its subordination to the methods and results of the sciences." If a naturalist position means a necessary relationship between philosophy and the natural sciences, we

Nietzsche's philosophy according to some perspectives that show the contributions and limits of the association between philosophical naturalism and the question of morals.

In addition to the reasons cited, the motivation for this study arose from the works of Brian Leiter, Christopher Janaway and Richard Schacht, which advocate that Nietzsche's philosophy is in consonance with philosophical naturalism. In face of the classical interpretations of Nietzsche's philosophy and the novelty that this striking affirmation represented, the central question of this paper is whether Nietzsche is a naturalist philosopher or not. A second question, originated from the first, is, if Nietzsche is indeed a naturalist philosopher, what is the contribution of naturalism to the problem of morals in his philosophy?

The most radical affirmation concerning philosophical naturalism in Nietzsche comes from Brian Leiter.³ For him, Nietzsche is a scientific naturalistic philosopher, and his naturalism is manifested in two ways: methodologically (which means affirming that Nietzsche is convinced that philosophy should be continuous with the results and methods of natural sciences) and substantively (which means affirming that Nietzsche believes that all existing things are natural). Brian Leiter believes that Nietzsche is actually committed to a scientific view about how things function. Brian Leiter is criticized by Christopher Janaway and Richard Schacht for relating the orientation of Nietzsche's thought with an exaggerated empirical scientism. Christopher Janaway⁴ advocates a wide version of naturalism in Nietzsche, supported by the negation of any kind of transcendental metaphysics and by the emphasis in the human body, in the instincts and the affections, in psychology and history (the disagreement seems to be mainly between a reductionist physicalist proposition and a non-reductionist physicalist proposition). Richard Schacht⁵ is even more distant from Brian Leiter's proposition and advocates an extended naturalism in Nietzsche. He affirms that Nietzsche's philosophy cannot be related to any type of methodological or metaphysical naturalism since the investigative methods employed by the philosopher are not in continuity with

should question whether this association is something required in some circles in contemporary philosophy—which desire to produce philosophical knowledge from this type of explanation—or if this association between philosophy and the natural sciences is part of Nietzsche's personal interest" (LOPES, Rogério. "A ambicionada assimilação do materialismo: Nietzsche e o debate naturalista na filosofia alemã da segunda metade do século XIX". In: *Cadernos Nietzsche* 29, p. 309-352, 2011, p. 311-312).

³ LEITER, Brian. *Nietzsche: on morality*. London: Routledge, 2002; LEITER, Brian. "Nietzsche's naturalism Reconsidered". University of Chicago, In: *Oxford Handbook of Nietzsche*, 2009, n. 35.

⁴ JANAWAY, Christopher. *Beyond Selflessness. Reading Nietzsche's Genealogy*. Oxford: Oxford University Press, 2007; GEMES, Ken & JANAWAY, Christopher. "Naturalism and Value in Nietzsche." In: *Philosophy and Phenomenological Research* 71, 729-740, 2005.

⁵ SCHACHT, Richard. "Nietzsche's Anti-Scientific Naturalism" In: Helmut Heit, Günter Abel & Marco Brusotti (Eds.): *Nietzsches Wissenschaftsphilosophie. Hintergründe, Wirkungen und Aktualität*. Berlin; New York: de Gruyter, 161-186, 2012; SCHACHT, Richard. "O naturalismo de Nietzsche". In: *Cadernos Nietzsche*, n. 29, p. 35-75, 2011.

scientific methods. In addition, the observations concerning human morals, even if related to physiology and psychology, are constantly molded by culture. Therefore, if there is naturalism in Nietzsche, it is constituted from a new sense that includes in the notion of natural a diversity of physiological, psychological, scientific, social, cultural, and artistic perspectives that are equally associated in the interpretation of the main philosophical problems.

From these arguments, it is necessary to assess whether or not a naturalistic statute (the necessary relationship between philosophy and the natural sciences) is demanded only by contemporary philosophy, that is, a contemporary reclamation of philosophy far from the time when Nietzsche wrote his texts and established his main philosophical problems. In face of the difficulty of supporting any type of *closed* philosophical naturalism in relation to Nietzsche's thought, it is necessary to discuss whether or not the association of Nietzsche's philosophy with naturalism is more plausible to his interpreters (focused on a debate that aims to make Nietzsche's philosophy "fit" into some type of naturalism, making the necessary conceptual stretches in order for that to be possible) than to Nietzsche's own philosophical propositions.⁶ In addition, Richard Schacht's analyses and Helmut Heit's reference indicate that the meaning of nature to Nietzsche seems to be distant from the meaning of nature used in a rigid naturalistic debate, in which nature means exclusively the material, the physical, that which can be explained by the laws of physics.⁷ Thus, to understand Nietzsche's intention of "re-translating human being back into nature,"⁸ it is necessary to investigate the meaning of nature to Nietzsche in association with his proposition of *re-translation*, marked by the potentialities of interpretation and the project of substitution of the vocabulary inherited from the traditional Western metaphysics in the presentation of the problems of morals and philosophy.

In addition, the main naturalistic projects of modernity should be analyzed as well as the perspectives and works that were known and read by Nietzsche and how they have influenced his views. In this approach, an informative pathway of reconstruction of Nietzsche's historical context is proposed by the studies by Thomas Brobjer⁹ and Rogério Lopes,¹⁰ which are interested in presenting Nietzsche's main interlocutors in modernity. Therefrom we know that Nietzsche was interested in learning the rigor that can be offered by the discipline and method of the natural sciences to the philosopher in the pursuit of philosophical conclusions, which is different from the exclusive fondness of natural sciences' methods and results advocated by Brian Leiter.

⁶ LOPES, Rogério. "A ambicionada assimilação do materialismo: Nietzsche e o debate naturalista na filosofia alemã da segunda metade do século XIX". In: *Cadernos Nietzsche* 29, p. 309-352, 2011.

⁷ HEIT, Helmut. "Perspectivas naturalizantes de Nietzsche em *Além do bem e do mal*". Transl. Leonardo Camacho. Accepted for publication in October 2014, forthcoming in: *Estudos Nietzsche*, v.6, n.1, jan./jun. 2015.

⁸ JGB/BM §230.

⁹ BROBJER, Thomas. *Nietzsche's philosophical context: an intellectual biography*. Illinois: University of Illinois Press, 2008.

¹⁰ Idem.

In face of Nietzsche's rejection of any type of evolutionary determinism and the criticisms to the notion of struggle for existence, self-conservation and vital competition, it is still necessary to investigate the possibilities of any naturalistic influence on the notion of will to power, which could have been inherited from the works of Darwin and the biologists Wilhelm Roux and William Henry Rolph, as identified by Wilson Frezzatti Júnior.¹¹ In this case, if there is any influence, it is not explicit, and the notion of will to power can only be considered the groundwork of a project of naturalization of morals if this does not mean a reduction of culture to biology or a type of biologism that affirms the biological origins of culture.¹² Resuming the possibilities of naturalism in Nietzsche's philosophy, it could only be configured as a new type of naturalism (still undefined in contemporary philosophy) that would not accept any type of reductionism and could encompass all the definitions of what is understood as culture.

Thus, the will to power can be interpreted in many different ways, but it seems accurate to say that this notion plays a fundamental role in the Nietzschean project of overcoming the metaphysical foundations of the soul and of opening the possibilities for new interpretations of the meaning of soul.¹³ From this argumentation, it is necessary to resume classic interpretations and readings of Nietzsche's works, which are responsible for giving his thought the relevance and importance that it currently has. This is an imperious need since the hermeneutics of the fundamental problems in Nietzsche's thinking has already been undertaken with competence and precision by his first interpreters and somehow the exaggerated preoccupation in finding the proper type of naturalism to his thought may have neglected these interpretations, which had already focused on and perceived its naturalistic nuances (and may contribute the groundwork for refuting some exaggerations). This far-fetched interpretation toward naturalism seems to occur with the notion of will to power as in the relationship hastily established between the will to power and a monist materialistic reading. Thus, against an excessively reductionist reading of the will to power, the interpretation of Wolfgang Müller-Lauter,¹⁴ as well as his arguments distancing Nietzsche's philosophy from Heidegger's readings¹⁵ should be revisited. Müller-Lauter's interpretation dissociates the will to power from a causal confinement within the physical world, avoiding a reductionist physicalist naturalism, and also from a metaphysical and cosmological

¹¹ FREZZATTI JÚNIOR, Wilson Antônio. "Tradução dos póstumos de Nietzsche sobre Darwin". In: *Estudos Nietzsche*, v.1, n.2, Curitiba, p. 403-419, jul./dez. 2010; FREZZATTI JÚNIOR, Wilson Antônio. *Nietzsche contra Darwin*. São Paulo: Discurso; Ijuí: Ed. UNIJUÍ, 2001.

¹² FREZZATTI JÚNIOR, Wilson Antônio. *A fisiologia de Nietzsche: a superação da dualidade cultural/biologia*. Ijuí: Ed. UNIJUÍ, 2006.

¹³ JGB/BM §12.

¹⁴ MÜLLER-LAUTER, Wolfgang. *A doutrina da vontade de poder em Nietzsche*. Transl. Oswaldo Giacoia Júnior. São Paulo: Annablume, 1997.

¹⁵ HEIDEGGER, Martin. *Nietzsche*. Vol. I e II. Transl. Marco Antônio Casanova. Rio de Janeiro: Forense Universitária, 2007.

aprioristic confinement, demonstrating the need of supporting some connection with empirical experience.

Taking Müller-Lauter as a starting point, we emphasize Nietzsche's commitment to knowing the method and results of natural sciences, but we also perceive that any form of causal, and even monocausal, explicative model is not admitted in his philosophy. Rogério Lopes stresses this argument in affirming that, if there is any naturalism in Nietzsche, it is a liberal type of naturalism that endorses the submission to an empirical discipline and a temporary scientific practice as a way of learning excellent conditions and the epistemic virtues necessary for philosophical reflection, with the aim of defeating the epistemic vices inherited from a metaphysical tradition. In this sense, a liberal naturalism advocates this experience by acknowledging the descriptive competence of natural sciences resulting from the sobriety and economy of the hypotheses demanded by the scientific method. The learning of this economy should be reflected in the genealogical, normative and terminological propositions, which could be perceived in the notion of will to power engendered in a wide sense of interpretation.¹⁶ In addition, in a liberal naturalism, suspicion should be the main characteristic of the philosopher in order to produce new ways of interpretation as a form of reaction to any form of dogmatism that intends to establish itself in thought to acknowledge the importance of values other than the cognitive and epistemic ones for philosophical reflection.

The study of Nietzsche's relationship with the natural sciences shows the philosopher's refusal of any form of determinism and causality as well as the interest in overcoming Western traditional metaphysics. Nietzsche provides many criticisms to the foundational solutions of morals. Thus, if Nietzsche, from his project of transvaluation of all values, is interested in destroying metaphysics and overcoming foundational theories of morals, how could the relationship of naturalism with moral normativity be conceived of? Jesse Prinz¹⁷ affirms that Nietzsche's interest in descriptive questions concerning morals is relevant since he is searching, with this question, for the origin of our moral values. This aim is evidenced by a genealogical intention of interpreting moral normativity, hoping that the exposition of the origin of values from our social history is able to destabilize established beliefs and function as a motivation to overcome these beliefs (in perceiving that the values prized by us have an origin, and this origin may not be that interesting).

Jesse Prinz seems to be right in relating Nietzsche's genealogy to the possibility of moral revisionism; nevertheless, he seems to disregard the importance that the notions of sense, value and history have within a genealogical reflection. Thus, it is important to show that, for the genealogical method to be effective as a strategy of moral revisionism, it is crucial to resume the interpretations provided by Gilles Deleuze¹⁸ and Michel Foucault¹⁹

¹⁶ JGB/BM §13.

¹⁷ PRINZ, Jesse. *The emotional construction of morals*. New York: Oxford University Press, 2007.

¹⁸ DELEUZE, Gilles. *Nietzsche e a filosofia*. Trad. António M. Magalhães. Porto: Rés, 2001.

concerning genealogy. To Gilles Deleuze, Nietzsche's critique of truths and metaphysics is necessary to withdraw from substantialism in philosophy and re-introduce the importance of interpretation. To Michel Foucault, the connection between genealogy and historical sense is necessary in order not to be involved by any supra-historical sense and, thus, aided by an effective history, re-introduce in becoming and corporeity everything that had been understood as immortal in the human being.

This does not mean that it is not possible to relate Nietzsche's project to some type of moral normativity. On the contrary, his previous arguments intend to show the pathways through which his thought can be connected to the notion of moral normativity without falsifying his philosophical intentions. Richard Schacht²⁰ affirms that internalization and identification are the pathways from which moral norms acquire normative effectiveness. However, in order to understand how morals develop and mainly how morals acquire a normative statute, it is necessary to investigate how internalization takes place in human beings. In this sense, it is relevant to resume Helmut Heit's affirmation about nature in Nietzsche, which sustains that the meaning of nature for the philosopher is in front of the meaning of nature used in the strict naturalistic debate. Thus, the Nietzschean task of "re-translating human being back into nature" is not related only to the task of destroying Western metaphysics to erect new notions of interpretation of the human phenomenon. Nietzsche intends to show that the effectiveness of the task of *re-translation* depends upon the reconstruction of the meaning of nature in Western philosophy to, therefrom, suggest how morals and culture arise within this reconstructed notion of nature.

The notion of nature cannot be harmonized with morals. For instance, for Nietzsche, the Stoic maxim that urges us to "live according to nature" points to an ideological conduct and a strategy to grant universal validity to a valuation. In Nietzsche's perspective, a well-grounded philosophy always imposes its configurations to nature. If avoiding it is not possible, Nietzsche recommends moderation in relation to believing in philosophy itself, which means believing that it is possible to use strong philosophical perspectives at the same time that its limited validity is acknowledged.²¹ According to Pierre Hadot,²² Nietzsche suggests that any knowledge about nature that does not result from anthropomorphism is inhuman knowledge. Thus, the solution would be to integrate into the interpretation of nature mistakes, illusions and the perception that values and representations are engendered by nature itself on behalf of life; they

¹⁹ FOUCAULT, Michel. *Nietzsche, a genealogia e a história*. In: Microfísica do poder. Translated by Roberto Machado. Rio de Janeiro: Edições Graal, 2012.

²⁰ SCHACHT, Richard. "Nietzsche's naturalism and normativity", in: JANAWAY, Christopher. (ed.); ROBERTSON, Simon. (ed.). *Nietzsche naturalism and normativity*. New York: Oxford University Press, 2012.

²¹ HEIT, Helmut. "Perspectivas naturalizantes de Nietzsche em *Além do bem e do mal*". Transl. Leonardo Camacho. Accepted for publication in October 2014, forthcoming in: *Estudos Nietzsche*, v.6, n.1, jan./jun. 2015.

²² HADOT, Pierre. *O véu de Ísis: ensaio sobre a história da ideia de natureza*. São Paulo: Loyola, 2006.

are not disposable and cannot stay out of any theory about nature. Just as the veils and the illusions are engendered by nature itself on behalf of life, nature's pure truth is the negation of life—the exaggerated will to truth becomes a will to death. The depth of things is horrifying, but appearance is engendered from this horror; nature creates representations and appearances to affirm existence.

If nature creates representations, nature is the quintessential artist. Therefore, Nietzsche posits that art has not a direct relationship with the beaux arts; the meaning of art encompasses all the activities of production and creation linking life to nature. Art is a force of nature that engenders the possibility of living. "Whoever has seen deeply into the world has doubtless divined what wisdom there is in the fact that men are superficial,"²³ says Nietzsche. This citation from *Beyond Good and Evil* has some similarity with the theses of his early texts about Greek mythology and arts. In an interpretation that overflows the objectivity of Heraclitus's fragment, Nietzsche could affirm that "Nature (Truth) loves to hide, loves to lie, loves illusion, loves to create works of art," as emphasized by Pierre Hadot with the aim of showing that Nietzsche does not endorse a strictly scientific interpretation of nature, demonstrating the value of a psychological and aesthetical approach in its interpretation.

Moreover, in analyzing Nietzsche's philosophy, it is important to perceive nature as a process in which the states and phenomena are engendered from the inter-relationship between consciousness and language. Thus, if in Nietzsche's thinking nature is a process, it should be thought of beyond dualistic and monistic propositions. Nietzsche admits some continuity between inorganic, organic, mental states, consciousness, becoming conscious, cognitive activities and the projects of action and its engenderments. As affirmed by Günter Abel,²⁴ to Nietzsche, human being is an embodiment of all the oldest estimations of value interpreted by all intelligent activities that have been found in the organic, became organic and from this moment on have participated in the organization of experience. To resume the issue of internatilization, we can say that Nietzsche affirms that "the will to power interprets"; therefore, appropriation and incorporation mean a will to configure and reconfigure until the subdued force has passed entirely to the power of the attacker and has potentialized it; if this incorporation is not successful, the formation will probably disintegrate.²⁵ Nietzsche considers that it is possible to think, feel and want without all these operations entering into consciousness since the preconscious and organic forms of life represent dynamic processes with the ability to interpret. The impulses are not blind—they are constantly occupied in observing, perceiving, ordering, delimiting, intensifying, preferring, despising.²⁶

Thus, according to Günter Abel, Nietzsche's critique to consciousness, to becoming conscious and to self-consciousness means a critique to the Cartesian

²³ JGB/BM §59.

²⁴ ABEL, Günter. "Consciência, linguagem e natureza: a filosofia da mente em Nietzsche". In: *Nietzsche na Alemanha*. Trad. Clademir Luís Araldi. São Paulo: Discurso Editorial; Ijuí, RS: Editora Unijuí, 2005.

²⁵ NF/FP 9[151] 1886-1887.

²⁶ JGB/BM §36.

model of consciousness, which affirms the possibility of a pure internal experience of the conscious being. To Nietzsche, we have consciousness and thinking due to language, signs and the need of interpretation, not through them, that is, it is not possible to think outside linguistic signs and, when attempting to do so, thinking is disrupted. Consciousness and language are engendered together. Human beings are capable of inventing signs of communication—these processes are linked to sociability and, the more they create signs of communication, the more they become conscious.²⁷ To understand the notion of nature in Nietzsche, it is necessary to know his interpretation of Heraclitus's fragment "nature loves to hide," his interpretation of ancient Greek art and culture and how he develops a philosophy of the process that connects nature, consciousness, language and sociability.

Similarly, to understand the project of transvaluating values, it is necessary to perceive the importance that psychology has in Nietzsche's thought, the motives that led him to affirm that psychology is the "queen of the sciences, for whose service and equipment the other sciences exist," that "psychology is once more the path to the fundamental problems"²⁸ and why in *Ecce Homo* he deems himself "a psychologist who has not his peers."²⁹ In his early writings, during the time when he taught in Basel, Nietzsche already posited that "philosophy develops its science of nature around false psychological data and involves the whole with a metaphysical necessity."³⁰ For him, the inside (psychological) has a correlation with the outside (physiological); therefore, he wants to free psychology from metaphysical prejudices to, from psychology, analyze moral prejudices. As affirmed by Giacoia,³¹ in order for that to be possible, it is necessary to perceive the objective of his critical philosophical project, which intends to destroy the metaphysical pillars of rational psychology and the theoretical bases of psychology in general to found a science of subjectivity devoid of metaphysical and moral prejudices that have founded Western thought, affirming the psychic connected with consciousness and unconsciousness. In order to destroy psychology's moralistic background, it is necessary to dissolve the superiority of consciousness and the idea of subjectivity identified with it. In addition, freeing psychology from moral prejudices may enable the liberation from nihilism as psychological state, overcoming the search for meaning, a moral canon, a universal order, a totality and an objective that is not in the events. The cause of nihilism is the dogmatic belief in the categories of reason because we learn to measure the world according to these categories that result from a fictive world. The new psychology intended by Nietzsche should demonstrate that the origin of moral values in modernity is the "thou shalt," that is, submission. The mark of modern morals is a combination of Platonism with Christianity; the systems of moral

²⁷ FW/GC §354.

²⁸ JGB/BM §23.

²⁹ EH/EH, *Por que escrevo livros tão bons*, §5.

³⁰ NF/FP 22[107] 1877.

³¹ GIACOIA JÚNIOR, Oswaldo. *Nietzsche como psicólogo*. São Leopoldo: Unisinos, 2006.

evaluation and judgment are derived from this combination, and psychology may be the science that will offer a way to overcome this morals.

The interpretation of nature in Nietzsche, differentiating itself both from Western metaphysics and a scientific view of nature, aims to build a perspective that connects the organic, the inorganic, consciousness and language. In this purpose, psychology may mean a science of depth, showing how nature engenders itself in organisms, moral values, societies and States. In Nietzsche's philosophy, psychology is not only a science of consciousness, but a science capable of revealing the depth that is veiled in human consciousness, the processes and the artifices of engenderment in nature. In this sense, the world is a self-generating and self-conserving organism, natural life a constant flow, and there is no such thing as an essential constitution. The forms of development of life can be investigated by perceiving that from the interplay between many wills to power, different engenderments may come about. This means that will itself is an illusion since this is one of the ways through which nature engenders itself—"creating" illusions. Thus, Nietzsche affirms that illusions are always more elaborated than nature. However, as affirmed by Araldi,³² Nietzsche understands human beings as units of force of relative duration and demonstration of stability, which are able to produce arts of transfiguration of unbridled violence and cruelty operating between the will to elude and the will to be eluded. Yet, the will to illusion does not arise from a human freedom to mold oneself, but the permanent impetus and pressure of a creative force.

These creative forces are related to dispute (agon) and with a disquieting double character of nature. Nature is disquieting because it contains an awesome aspect that it struggles to hide. Thus, a tragic wisdom should be engendered that acknowledges the impossibility of rationalizing and knowing nature as a whole, which is backed by the knowledge of the limited character of its institutions, standards and norms. In this notion of nature, the understanding of the limits of knowledge, the impossibility of the will to truth and the perception that nature is awesome are demanded, but we cannot live without this knowledge exactly because it shows the limits of the possibility of knowing. Nature is engendered from the possibility of dispute and conflict. The continuity of a society and a State lies in its capacity to be in favor of dispute and to distance itself from conflict. To Nietzsche, the prolonged maintenance of the absence of war represents one of the most interesting aspects produced by dispute, but this is in relation to the capacity of this society or State to grasp the tragic element in its culture.³³

Is Nietzsche a naturalistic philosopher? He could only be considered so if the type of naturalism related to his thinking is also related to the notion of nature. In this case, this naturalism is a new naturalistic approach, different from the

³² ARALDI, Clademir. "Organismo e arte na filosofia de Nietzsche". In: *Philosophica*, 29, Lisboa, 2007.

³³ NIETZSCHE, Friedrich. "A disputa de Homero". In: *Cinco prefácios para cinco livros não escritos*. Trad. Pedro Sússekind. Rio de Janeiro: 7 letras, 2007; NIETZSCHE, Friedrich. "O Estado grego". In: *Cinco prefácios para cinco livros não escritos*. Trad. Pedro Sússekind. Rio de Janeiro: 7 letras, 2007; NIETZSCHE, Friedrich. *O Nascimento da Tragédia: ou Helenismo e Pessimismo*. Transl. J. Guinsburg. São Paulo: Companhia das Letras, 1992.

naturalistic categories that exist in contemporary philosophy and, at the same time, the notion of nature that grounds this new type of naturalism is distanced from the notion of nature understood in contemporary naturalism. However, naturalistic perspectives can be found in his philosophy, which was perceived by his main commentators, and his naturalistic perspectives are connected to his project of transvaluation of values.

THE NATURALIZATION OF MORALS AND THE FAREWELL TO NATURALISM IN NIETZSCHE'S PHILOSOPHY

CLADEMIR LUÍS ARALDI

Despite the many naturalistic approaches to Nietzsche's philosophy that emerged in the last decades,¹ defining Nietzsche's naturalistic position in relation to morals is still a difficult task. For such, it is necessary to rebuild the strong connection between naturalism and genealogy concerning moral values. The naturalistic explanation to the origin of moral values, as I intend to show, is in consonance with the genealogical method in its three main axes: psychology, history and physiology. Thus, this study will investigate naturalism in the sense formulated by Nietzsche firstly in *The Gay Science*. At this point, it is important to resume Brian Leiter's questioning² about whether Nietzsche's naturalism is an instrument for the "revaluation of all values."

The difficulty consists exactly in articulating the two very different tasks of this project of Naturalization (*Vernatürlichung*): 1) the development of the critique of morals (including the dehumanization of nature), and 2) the naturalization of morals and man, since Nietzsche himself seems to tell the critique of morals from the invention of new values, as pointed out in the critical evaluation of all moral values of *On the Genealogy of Morals*:

we need a *critique* of moral values, *the value of these values should itself, for once, be examined* – and so we need to know about the conditions and circumstances under which the values grew up, developed and changed [...], since we have neither had this knowledge up till now nor even desired it.³

However, Nietzsche does not separate the critical from the affirmative aspects in developing the genealogy of moral values. Therefore, in the building of his "affirmative ethics," Nietzsche easily transposes the limits of the critical genealogy to accomplish his creative "task" of naturalizing moral values as if there was a necessary link between these domains. I intend to show that there is not a necessary connection between the critique and the invention of moral

¹ I highlight the approaches of B. Leiter (2002), R. Schacht (1983,1994), Ch. Janaway (2007).

² See LEITER, *Nietzsche on Morality*. New York: Routledge, 2002, p. 26: "Nietzsche develops a naturalistic account of morality in the service of a very particular *normative* goal, namely, to force us to reconsider the *value* of morality: naturalism is enlisted in the service of what Nietzsche calls his "revaluation of all values"."

³ GM, Preface, § 6.

values in the works *The Gay Science* (GS), *Beyond Good and Evil* (BGE) and *On the Genealogy of Morals* (GM), but an urgency to overcome moral nihilism. In this sense, this study is articulated into three points: 1) The dehumanization of nature (*Die Entmenschlichung der Natur*) through the moral critique in *The Gay Science*; 2) the naturalization of man (*die Vernatürlichung des Menschen*) and their values in GS, BGE and GM and 3) the central role played by art in the creation of naturalistic values.

I. THE DEHUMANIZATION OF NATURE (THE CRITICAL TASK)

The main point of Nietzsche's naturalism is establishing the relationship between philosophy and the sciences from the (pre-)genealogical method. In the core of these naturalistic investigations lie moral values. Here, the scope is limited to the relationship that can be made in Nietzsche's philosophy with the "best" sciences of his time: physiology, psychology and history. The works in which the philosophy of the free spirit is developed – *Human, all too Human*, *Daybreak* and *The Gay Science* – are pre-genealogical insofar as they do not yet methodologically define the criterion for evaluating all moral values. However, they actually point out to the objective of Nietzsche's naturalism, namely that of uniting the ethical-philosophical research and the sciences.

After the publication of *Daybreak*, as he prepared *The Gay Science*, Nietzsche studied scientific works, mainly those by Robert Mayer, Wilhelm Roux, "*Der Kampf der Theile Im Organismus. Ein Beitrag Zur Vervollständigung der Mechanischen Zweckmässigkeitslehre*," as shown by W. Müller-Lauter, Günter Abel and Gregory Moore⁴.

Naturalism is formulated more clearly in the Book III of *The Gay Science*, after the first mention to the death of God (GS, § 108). Only after the utter *de-divinization of nature* the question of the *naturalization of man* would be put forth:

When will all these shadows of god no longer darken us? When will we have completely de-deified nature? When may we begin to *naturalize* humanity with a pure, newly discovered, newly redeemed nature!⁵

The connection between the critical aspects and the affirmative project also appear in a preparatory writing to the GS: "My task: the dehumanization of nature and, then, the naturalization of man after he has obtained the pure concept of 'nature'."⁶ The first task consists in describing nature without anthropomorphisms and without moral, aesthetic, and religious interpretations with which man has falsified and covered it throughout history. This would enable achieving, secondly, *the naturalization of man*. This is not a task for natural science strictly speaking, but for the *gaya scienza*.

In the Book IV of *The Gay Science*, however, Nietzsche seems to fundamentally modify the project of achieving "pure" nature, with no moral

⁴ Cf. G. ABEL, 1998, W. MÜLLER-LAUTER, 1971 and G. MOORE, 2002.

⁵ GS, § 109.

⁶ PF 11 [211] – beginning of 1881 – summer of 1882.

interpretation. In GS 335, with the title of “Long Life Physics,” he no longer develops the general character analysis of the world as chaos. The exhortation to physics takes place within the ambit of psychological observation and the emergence of moral judgments. To be creators of new values, “... we must become the best students and discoverers of everything lawful and necessary in the world: we must become *physicists* in order to be creators in this sense – while hitherto all valuations and ideals have been built on *ignorance* of physics or in *contradiction* to it.”

This is the necessary condition not only for *new naturalistic values*, but also to configure and *create oneself*. The emphasis lies in the ethical-aesthetical aspect, which, strictly speaking, surpasses the limits of the natural sciences. According to Nietzsche, however, man always projects valuation perspectives in natural sciences, including physiology.

By that time, Nietzsche intends to effect “the reduction of morals to aesthetics.” The task involves creating “a profusion of aesthetic valuations, equally justified [...], each of them being the ultimate fact and the measure of things for the individual.”⁷ Nietzsche’s “Naturalism” is above all an “experiment of the knower,” from which he wants to configure a *gai saber*. It is the rare art of “giving style to its character” (GS 290), so that man may be the “poet of his own life” (GS 299). This means that both scientific praxis and the creation of new values would have an aesthetic character.

We question, however, whether by the time GS was written Nietzsche was already a speculative methodological naturalist, as claimed by Brian Leiter in his 2002 book *Nietzsche on Morality*. Every phenomenon has deterministic causes so that they can be naturalistically understood. The deterministic explanation of how men act, feel and think the way they do would have been employed by Nietzsche within ethics.⁸ According to Leiter, however, in his late work (especially BGE and GM), Nietzsche provides deterministic causal explanations for the human phenomena, mainly within ethics. The positive point is the continuity of methods with the empirical sciences. But would that be the decisive naturalistic turn in *Beyond Good and Evil*? For B. Leiter, naturalism is no more than a part of Nietzsche’s philosophical project, the only, however, that would be of relevance for ethics and moral philosophy.

Nietzsche would be reductionist if he affirmed that values were only direct expressions of physiological impulses and needs. This occurs in Leiter’s naturalistic explanation of moral values and beliefs from the individual psychophysical constitution (the Typ-facts). To avoid this reductionism, Christopher Janaway questions the coherence of *methodological naturalism* in Nietzsche’s thinking. Janaway advocates that there are no scientific proofs (both in relation to the methods and in relation to the results of the science) for many of Nietzsche’s explicative hypothesis, such as, for instance, the introduction of

⁷PF, 11[79].

⁸ Cf. LEITER, B. *Nietzsche on Morality*, Routledge, New York 2002.

the will to power as the “best model” for biology and the organic world.⁹ In addition, Nietzsche, as a naturalist, would not completely employ the scientific methods for understanding things. If Nietzsche is a naturalist since *The Gay Science*, in what sense can we speak of naturalism?

Nietzsche did not simply find in the sciences of his own time the naturalistic criterion to evaluate moral values. His insights and his “deep thoughts” surpass the empirical domain of science, just as the conceptual-argumentative ambit of philosophy. However, he searched for and found in the sciences some methods and results to prove them, since he believed they could provide a scientific character to his thinking.

II. THE NATURALIZATION OF MEN AND THEIR MORALS

The project of naturalization of man has a significant elaboration in *Beyond Good and Evil* and in *The Gay Science*, Book V, when Nietzsche develops the natural history of morals. We can affirm that, in a certain sense, from this time on Nietzsche becomes an *ethical naturalist*. With this definition, emerge also the difficulties of understanding the singular character of his enterprise. The challenge consists in maintaining a naturalized conception of Nietzsche’s genealogy without abandoning the positive and critical tasks of his late philosophy in relation to ethics. My thesis is that in his naturalistic interpretations, Nietzsche goes beyond the genealogical critique of morals without determining his ethical position.

The genealogist of morals assumes a modest empirical task: formulating and conceptually classifying the different moral experiences of value, the long and difficult task of understanding “the hieroglyphic writing of human moral past” (GM, *Preface*, § 7). In this context he operates with physiology as a starting point to his genealogical-naturalistic researches.¹⁰ It is also physiology that allows articulating naturalism and genealogy. We cannot forget that Nietzsche necessarily links physiology with psychology and history. According to my interpretative hypothesis, the genealogical method will be effective (among the three methods mentioned above) only if the philosopher shows that the will to power is the explicative model or a valid criterion for the naturalization of morals, that is, to assess moral values in a non-moral way. In this sense, the trial of naturalization of morals depends on this impulse to power in its pshysio-psychological aspects and in its manifestations throughout history.

Thus, one attempts to understand Nietzsche’s particular naturalism as a physio-psychological hypothesis that aims at explaining the human ways of feeling, thinking and evaluating, above all in relation to the emergence and

⁹ “On a straightforward reading, Nietzsche goes out of his way to reject Results Continuity with scientific biology – unless he believes that a perfected scientific inquiry would find that relations of over-powering and interpretation were indeed the best models for biological process. But in that case more recent science does not display Results Continuity with Nietzsche.” (Christopher Janaway, *Naturalism and Genealogy*, In Keith Ansell-Pearson, *A companion to Nietzsche*, Blackwell 2006, p. 340).

¹⁰ Cf. PF 40[21], outono 1884 – outono de 1885.

transformation of human values. I question here whether Nietzsche develops naturalism strictly speaking, as he announces in the end of his first dissertation in GM, namely that genealogists should prepare the future task for philosophers: that of creating values. If the task of genealogy articulated to naturalism is thus exhausted, the Philosopher Nietzsche tries to abandon naturalism even before having developed the potentialities of this method.

Philosophers had up to now only an insufficient knowledge of physiology, says Nietzsche in the beginning of BGE. Therefore, it would be necessary to free the physiology of his time of moral prejudices and superfluous teleological principles, such as the principle of self-conservation.¹¹ On the other hand, the critic of metaphysics wants “to practice physiology in good conscience,”¹² duly treating physiological demands and physiological constitution.¹³ Thus, it would be possible to advance in the constitution of an “authentic physio-psychology,” that is, the combination of physiology and psychology. This physio-psychology (*Physio-Psychologie*), as a “morphology and theory of the development of the will to power,”¹⁴ would be the most promising path to the naturalization of morals. The physiopsychologist would have the prerogative to duly interpret nature, which is the evaluator and dominator of human impulses, beyond moral prejudices. Nobody had gone this way until then, and Nietzsche does not advance much in it.

Moral is the greatest impediment to resuming the *homo natura*. So the insane task is

To translate humanity back into nature; to gain control of the many vain and fanciful interpretations and incidental meanings that have been scribbled and drawn over that eternal basic text of *homo natura* so far; to make sure that, from now on, the human being will stand before the human being, just as he already stands before the *rest* of nature today, hardened by the discipline of science [...].¹⁵

Moral prejudices would have prevented the investigation of physiological processes and facts that are the groundwork for all the judgments and values of good and evil. The repeated mentions to the physiology of power (*Physiologie der Macht*), the physiology of passions (*Physiologie der Leidenschaften*), the physiology of morals (*Physiologie der Moral*) or to the relationship between morals and physiology occur in the sense of reinforcing this investigation.¹⁶ Physiology undoubtedly has a decisive role in the first dissertation of the GM in criticizing the failures (*die Missrathenen*) and for the position of values of the noble type of man. Nietzsche describes “physiological facts” (*physiologische*

¹¹ Cf. BGE, § 13.

¹² Cf. BGE, § 15.

¹³ Cf. BGE, § 3.

¹⁴ Cf. BGE, § 23.

¹⁵ BGE, § 230.

¹⁶ Cf. *PF* 27[14] and 27[37], summer – fall of 1884; *PF*, 29[67]; *PF* 37[4], June – July of 1885; *PF* 1[186], e *PF*2 [76], fall of 1885 – fall of 1886.

Tatsachen) concerning human nature, attempting to explain how Christian values came out of the resentment of failures.

In concluding his first dissertation with the narration of the terrible millenary struggle for power between the values of slave morality and master morality, Nietzsche places himself in a perspective that goes beyond good and evil as conceived in the old morals. But this does not mean that he is free from valuating positions, since the aristocratic way of valuing is the perspective he adopts in his critical and re-evaluating work.

It is meaningful that Nietzsche was aware of the immensity of the task of naturalizing genealogy. In admitting the limits of his philosophical enterprise, the solitary philosopher expresses a wish in the note that closes the first dissertation of GM: “it is just as essential to win the support of physiologists and doctors for these problems (on the *value* of all previous valuations).”¹⁷ The professional philosophers would be no more than mediators to render the relationship between physiology, medicine and philosophy fruitful. In this “scientific interchange,” physiology seems to have a greater significance among the sciences, since it helps the philosopher to solve the problem of value; it is not, however, developed in physio-psychology as announced in BGE I, 23.

Up to this point (at the end of GM I), Nietzsche is neither a *naturalized genealogist* nor yet the *Philosopher* who would be in condition of solving this great problem. But the new relationship between sciences and philosophy presented there is promising: “All sciences must, from now on, prepare the way for the future work of the philosopher.”¹⁸ This “future Philosopher” would actually be a legislator and a creator of new values. However, he does not exist yet, although Nietzsche yearns for him:

– Towards *new philosophers*, there is no alternative; towards spirits who are strong and original enough to give impetus to opposed valuations and initiate a revaluation and reversal of “eternal values”; towards those sent out ahead; towards the men of the future who in the present tie the knots and gather the force that compels the will of millennia into *new* channels. To teach humanity its future as its *will*, as dependent on a human will, to prepare for the great risk and wholesale attempt at breeding and cultivation and so to put an end to the gruesome rule of chance and nonsense that has passed for “history” so far [...] a new type of philosopher and commander will be needed for this some day.¹⁹

Is man still unexhausted for great possibilities? In case the answer is yes, then would we all, herd animals, be infected by the “complete degeneration of man” (*Gesamt-Entartungdes Menschen*), that is, by the process of nihilistic dissolution? Were we irreversibly dominated by the will to nothing? Or are the inexhaustible energies of the *homo natura* still found in us? This is Nietzsche’s question mark, which seems to us exaggerated preoccupations, typical of the end

¹⁷ GM I, § 17.

¹⁸ GM I, § 17.

¹⁹ BGE, 203.

of the 19th century. However, they bring to light the disquieting problem of the value of human values.

* * *

For our problem—the naturalization of morals—we arrive at a provisional conclusion: without the sciences of nature, the Philosopher cannot act. I understand that Nietzsche the *genealogist* assumes very partially the task of articulating philosophy with physiology, medicine, psychology, ethnology, philology, and history in his own time. Therein lies his surprising modesty: genealogists and future philosophers have the task of developing and consummating the naturalization of morals. Nietzsche himself does not go further with the project of naturalism in this strong sense. His artistic inclinations and the not clarified belief that art is the only antidote to nihilism lead him into following another path, saying farewell the naturalism associated to the natural sciences.

III. NATURALISM AND ART: THE RELATIONSHIP BETWEEN NATURALISM AND THE ART OF LIVING OF THE FREE SPIRIT

After the naturalistic investigation of the three psychological mechanisms (resentment, bad conscience and the pursuit of power in ascetic priesthood), Nietzsche affirms, in the end of GM, III (§ 25) that science, including modern science, is not yet the antagonist of moral denaturalization of the ascetic ideal. Science shares with the ascetic ideal “the same faith that truth *cannot* be assessed or criticized), and this makes them both *necessarily* allies.” Art, in which the will to mischief has good conscience, would be “more fundamentally opposed” to the ascetic ideal. In valuing art in this way, it seems to me that Nietzsche abandons the “heroism of veracity” and the epistemic virtues of the free spirit. There is no criterion yet of naturalistic evaluation, since the artistic inclination to simplification and falsification would be stronger than the artistic inclination to knowledge:

What the spirit enjoys here is its multiplicity of masks and its artfulness, and it also enjoys the feeling of security these provide, – after all, its Protean arts are the very things that protect and conceal it the best!²⁰

The good will to appearance through art is not the definitive solution in *On the Genealogy of Morals*. In searching for a more adequate solution, Nietzsche planned, in the fall of 1887, to write three more dissertations for the second part of the *Genealogy of Morals*. One would be dedicated to the study of “the history of denaturalization of morals.”²¹ By that time, he intended to replace “moral values with *naturalistic* values.”²² It is the continuation of the project of

²⁰ BGE, 230.

²¹ PF, 9 [83] – fall of 1887.

²² PF, 9 [8].

retroverting man to nature, which now is called “moralist naturalism,” namely “the retroversion of moral value apparently emancipated and supernatural to its ‘nature’, namely to its natural morality.”²³ Here we have a progress in Nietzsche’s naturalistic project insofar as it presents a new conception of art against the denaturalization of values: “I also want to naturalize Aesthetics.”²⁴ The development of the physiology of art shows that “ethics” is the objective of Nietzsche’s naturalism. Only through art, or better, only through the physiology of art it would be possible to create new values and ways of life to reach “a creative attitude towards the world and ourselves.”²⁵

Through the “naturalism of morals” (*Naturalismus der Moral*), Nietzsche aimed to describe the physiological realities “beyond morals.”²⁶ He employs the “will to ascending life” (*der Wille des aufsteigenden Leben*)²⁷ to criticize the symptoms of decline in the Christian morals of compassion and in Wagner’s music. In this sense, the real primary causes of moral and aesthetic values are physiological. In 1888, Nietzsche reinforces his position of 1887 (GM III) in affirming that art is more valuable than truth, since “art is the only force contrary and superior to the will of denying life.”²⁸ Here Nietzsche is not talking about modern art, which ends up in nihilism, in the exhaustion of value and meaning. The reduction of morals to aesthetics is a farewell to moral philosophy—and the entrance into a world devoid of values.

In face of the threat of nihilism, the solitary Philosopher ached to create naturalistic values: they are not “moral” values, since they have an aesthetic character and physiological presuppositions. Naturalizing aesthetics means that values are human inventions that arise from the will to mischief, illusion and appearance, but they can also lead to superior forms of life.

The supposed ethical-aesthetical Naturalist still searches for the physiological causes that result in the increase of power. The physiology of art could be a necessary means to naturalize values. We could then understand the reasons why Nietzsche resumes the themes of the Dionysian and inebriation in 1888. In inebriation an increase in force, a feeling of wholeness effectively occurs.²⁹ The investigation of the physiological causes of inebriation shows that this state may express two radically different types: i) a sick type, that masks its ill states (art would be an example for that) and ii) a healthy type, in which art expresses the affirmative instincts of life.

Unfortunately in the last months of his philosophical art of living, Nietzsche was more concerned with the ill aspects of inebriation, mainly the diagnosis of unnatural traits in morals. Maybe he had overestimated the danger of nihilism. Nevertheless, he opened a very promising pathway for the naturalization of

²³PF, 9 [86].

²⁴PF, 9 [93].

²⁵JANAWAY, C. *Beyond Selflessness: Reading Nietzsche’s Genealogy*. New York: Oxford University Press, 2007, p. 234.

²⁶ See PF, 16 [73] beginning – summer of 1888.

²⁷PF 16 [86].

²⁸PF, 17[3], May – June 1888

²⁹Cf. PF 1888, 17 [9].

morals, in which psychology, history and physiology could contribute for the future creator of the human being. For such, we should return to the philosophy of the free spirit.

In this direction, Giuliano Camponi's position about the "French Nietzsche" should be addressed. It is very valuable to fixate the *sui generis* character of Nietzsche's naturalism, which I call here *naturalisme à la provençale*. By the same time that he was searching for a new "physiological-psychic center," in which he was focused on the genealogical construction of the subject, Nietzsche believed he had found in Paris "the laboratory of values and forms of life," mainly in the budding empirical psychology. In addition to the readings of P. Bourge, Ch. Féré and Th. Ribot, he was focused on studying the new French *romanciers*.³⁰ He turns to the south, to the joyful and serene genius of the Provençal. The "France of taste" is not the "imbecile and rude France" of the bourgeois-democratic era in which he lives. The few who belonged there are hidden, obscure, sick. What mattered, however, is exactly the "inheritance" of the France of the spirit. In addition to the capacity for having artistic passions and the culture of the moralists, the superiority of the French of the good times was thus expressed:

At the core of the French there is a half-successful synthesis of north and south which lets them conceive many things and do many others that will never occur to an Englishman. Using a temperament that is turned periodically towards and away from the south, and whose Provençal and Ligurian blood bubbles over from time to time, the French fortify themselves against the awful northern gray on gray, the sunless concept-ghostliness and anemia, – our *German* disease of the taste, against whose excess people at the moment are strongly resolved to prescribe blood and iron: I mean "great politics"(following a dangerous medical practice that teaches me to wait and wait but not, so far, to hope –)³¹.

As a new free spirit (*freier Geist*), with new tasks, Nietzsche projected himself in his time as one of the rarest, who felt welcomed in France. As a "native Mediterranean," who learned to love the south in the north and the north in the south, he proposes the Philosopher's task: self-affirmation from an ethical-aesthetical perspective. As those who "determine the hierarchy of values,"³² the future philosophers would be legislators strictly speaking. They would be, above all, artists and legislators. Only in a small measure Nietzsche assumes a commitment with the epistemic virtues and with the methods of science. After all, it is only through art and artistic illusions that existence could be affirmed in face of the threats of nihilism. In this sense, the focus would become rather the naturalization of art (*die Vernatürlichung der Kunst*), with the support of a naturalized approach to asceticism.

³⁰Cf. Giuliano Campioni, *Der französische Nietzsche*. Berlin, de Gruyter, 2009, pp. 313 - 314.

³¹BGE, § 254.

³²GMI, § 17.

Asceticism of the spirit (*Der Asketismus des Geistes*)³³ is a condition for man to regain nature:

“I want to *renaturalize Ascetics*, too; instead of the tendencies to negation, the tendencies to *strengthening*; a gymnastics of the will; a deprivation and periods of voluntary fasting in all ways, also in the spiritual senses *auch im Geistigsten*.”³⁴

To reach self-domination, many men in different times had to obey during long periods, as well as practice many exercises of bodily and psychic deprivation, such as the ascetics of the Philosophy of Vedanta, the “practical ascetics” of the Greek philosophers or even the “more popular” ascetics of the old Germans.³⁵ The Free Spirit would also willingly engage in ascetic exercises. With discipline and rigor it would dedicate itself to the “gymnastics of the will,” that is, the asceticism of the spirit and of virtue, enjoying the overflowing feeling of power and plenitude in the solitary desert of its liberation. In the years of philosophical errands, the free spirit philosopher assumes the task, with patience and determination, to build the “asceticism of the strong” (*der Asketismus der Starken*).³⁶ The ascesis typical of these moral and bodily exercises would be a means to the intensification of power in order to reach the “innocence of the child in becoming”³⁷; finally, it is a means to give “style” to the character and existence of each and every one as it was proposed since *The Gay Science*.

³³ Nietzsche employed many expressions to mean “ascesis”: *Asketismus, Asceticismus, Asceticism, Asketism, Ascetik, Asketik, Asketenthum, Askese, Askesis, ascese*.

³⁴ PF 9[93] fall 1887

³⁵ Vgl. GM III, 12, M 195.

³⁶ Cf. PF 15[117] – spring de 1888.

³⁷ PF 25 [351]-Frühjahr 1884

**IS ONTIC STRUCTURAL REALISM A VIABLE FORM OF
NATURALIZED METAPHYSICS? NATURALISED
METAPHYSICS AND TROUBLES CAUSED
BY THE HISTORY OF SCIENCE**

JACK RITCHIE

Sometimes naturalism is presented as a general world view. Most often when so presented naturalism is identified with physicalism: the view that everything is physical or in some sense dependent upon the physical. But physicalism faces a well known and very serious problem, Hempel's dilemma. The problem can be put like this: when we say everything is physical what do we mean by the physical in physicalism. Do we mean to define the physical in terms of current physics? Well, then the view is almost certainly false. We expect current physics to be revised as science develops in unknown ways and so some of the claims currently accepted by physicists to be rejected as false in the future. Well, then do we mean some idealised future physics? Since we have no idea what that idealised future physics is, we have no real idea of what the content of physicalism is supposed to be on such definition. If we are told that physics here just means the discipline which captures all the facts which all other facts are dependent upon, then the claims of physicalism are tautological.

There is a vast literature on this problem but it is not my purpose to engage with that here. All we need to note is that Hempel's dilemma highlights a general problem, which a naturalist ought to be very sensitive, about combining naturalism with any very general world view. The history of science and physics teaches us that science is prone to revolutionary episodes. Basic ideas about how the world works are overthrown as physics develops. Once scientists thought, for example, there was absolute space, caloric and phlogiston but now we reject such ideas. So even modest reflection on science and its history ought to make us at the very least a little wary about drawing any confident conclusions about the fundamental furniture of the universe; and we would expect anyone calling themselves a naturalist to be at least that reflective on science and its history since for a naturalist it is from science she hopes to build her philosophy.

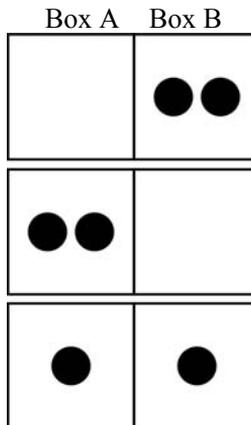
Does this mean that that naturalized metaphysics is impossible. Not necessarily. Perhaps metaphysicians can learn something from scientific realists. Presented with the a history of science of past failure realists typically claim that a more nuanced reading of the history can show that where there appears to be discontinuity and error, there is in fact gradual accretion and progress. One

metaphysical position which takes this approach is so called Ontic Structural Realism (OSR). In line with others that call themselves structuralist they believe that as science progresses we see structures retained from one theory to the next and so we have good reason to believe in those. To this epistemic thesis they add a radical metaphysical twist – all there is structure and, very importantly for this volume, a loud and confident proclamation that this is a radical form of naturalized metaphysics since its claims and arguments are drawn directly from science.

In the rest of the paper I explore this idea and its naturalist credentials. First I look at the argument which makes use of some odd aspects of contemporary physics to motivate the very radical thesis that there are no things. Then I consider structuralist responses to the problem of scientific revolutions before finally turning to consider whether the idea that structure is all there is makes any sense. My conclusions will be uniformly negative. The arguments for structuralism are poor and it is doubtful OSR makes sense.

OSR AS MOTIVATED BY PHYSICS

The argument for OSR from physics starts with some observations about the strangeness of quantum systems. Consider a system of two particles, illustrated below:



Classical physics says there are 4 possibilities here. Both particles can be in Box A, both in B, particle 1 in box A and particle 2 in B and vice versa .

In quantum theory different statistics apply depending on what kind of particle is under consideration. In particle physics we divide particles into two kinds - bosons and fermions. Bosons are particles like photons, and larger complex of particles like the hydrogen atom. They obey Bose-Einstein statistics. According

to these statistics there are only three possible states these particles could be in. There is only one state corresponding to the bottom line.¹

Fermions are particles like quarks and electrons. They obey Fermi-Dirac statistics. There is only one possible state the fermions could be in according these statistics, the one represented by the bottom line. What is especially odd about all of this, whether the particles we are interested in are fermions or bosons is that what we might think of as natural permutations of particles do not correspond to different physical states.

The fathers of quantum theory thought that this should lead to profound revision of our basic ontology. Here is how Hermann Weyl elegantly sums up the situation.

[T]he possibility that one of the identical twins Mike and Ike is in the quantum state E1 and the other in the quantum state E2 does not include two differentiable cases which are permuted on permuting Mike and Ike; it is impossible for either of these individuals to retain his identity so that one of them will always be able to say 'I'm Mike' and the other 'I'm Ike.' Even in principle one cannot demand an alibi of an electron! (Weyl 1931,)

Quantum particles they are argued could not be treated as individuals.

In the 1980s some philosophers of physics developed a different interpretation of the funny quantum statistics. They argued that quantum particles were individuals but whatever facts individuated the particles (some basic thisness or other metaphysical property)² were not represented in the theory. The particles obeyed these funny stats not then because of their lack of individuality but simply because of nomological constraints. Not all the states that one would expect classically to be available to a particle where in fact possible states. Here is how Steven French sums up this view:

[T]he implication of the different 'counting' in quantum statistics is not that the particles are non-individuals in some sense, but that there are different sets of states available to them, compared to the classical case. On this view, the particles can still be regarded as individuals — however their individuality is to be understood metaphysically (French 1989)

This leads to what advocates of OSR call a version of metaphysical underdetermination. One theory, quantum mechanics, can have multiple interpretations. So even if one believed that the empirical success of that theory were good grounds to believe it, metaphysical underdetermination would imply that you would have no way of working out what the content of theory is that

¹ So if appropriate symmetries are in place so all states are equally probable, then classically we would expect the probability that the particles are in different boxes to be $\frac{1}{2}$ but in Bose-Einstein stats the probability would be $\frac{1}{3}$.

² The Bohm theory would be an example of this. Here the hidden variable is the definite particle position.

you ought to believe. The success of the theory is compatible with multiple, contradictory metaphysical interpretations.

This sounds like very bad news for realism of any kind but James Ladyman and others have argued that in fact cases of metaphysical underdetermination like this should motivate OSR.

In the case of individuality, it has been shown [...] that electrons may be interpreted either as individuals or as non-individuals. We need to recognize the failure of our best theories to determine even the most fundamental ontological characteristic of the purported entities they feature. It is an ersatz form of realism that recommends belief in the existence of entities that have such an ambiguous metaphysical status. What is required is a shift to a different ontological basis altogether, one for which questions of individuality simply do not arise. (Ladyman 1998)

The argument is not elaborated much more than this in the publications of advocates of OSR but it seems to go as follows:

1. If we assume quantum objects, then our ontology is metaphysically underdetermined.
2. If our theories are metaphysically underdetermined, then we can't be realists about those theories since we do not know what we are being realists about.
3. If we adopt OSR, there is no metaphysical underdetermination.

Hence(?) if we want to be realists, our best bet is to be ontic structural realists.

Let's just assume 1 is right for the reasons given by advocates of OSR. The argument even so seems very odd for it arrives at a metaphysical conclusion by apparently appeal to an epistemic problem. We don't *know* which interpretation of quantum theory is right so we should move to a third interpretation – OSR. But on the face of it OSR makes the problem of metaphysical underdetermination worse. Instead of just having two underdetermined rivals, it looks like we now have three. Particles as individuals, particles as non-individuals or no particles just structures. Why should we favour any one of these options over the others?³

Steven French (2014) has recently argued that OSR is not just another competitor interpretation but picks out the “common core” between the underdetermined rivals and that is why we should favour OSR over the two rival interpretations. But if OSR were the common core between these rival interpretations, then it would not deny anything asserted by the other theories. But it does. That there are objects. So this claim is implausible. It could be argued that what is the common core is the mathematical structure which is common to both interpretations but even if we grant that is so no metaphysical conclusion follows. Surely the most sensible attitude to adopt would be one of epistemic modesty. We should draw back from assenting to any claims that go

³ This point is made in Saatsi (2009) and Brading and Skiles (2012).

beyond the structural. Such an attitude would be that of epistemic structural realism (ESR). ESR recommends that we believe in the structure of the theory but remain agnostic about all other claims. Hence in this case we ought to believe the fact about quantum systems contained in the group theoretic structures but be agnostic about whether there are particles and those particles are individuals. In short it seems to me that the favoured argument of metaphysical underdetermination offers no reasons to believe ontic structural realism but possibly some reason to accept ESR. With that I mind I know turn to the arguments from the history of science which are meant to motivate both positions.

OSR AS MOTIVATED BY THE HISTORY OF SCIENCE

As we noted above a fundamental difficulty for any metaphysical view with naturalistic pretensions is to offer some response to the problem of scientific revolutions. How can we get reliable information about how the world is from science given we expect our best scientific theories to be overthrown in the future? The general structure of any response to this problem will be to adopt a so called divide and conquer strategy. Scientific realists will typically argue that we have reasons to believe some parts of our theories but ought to be anti-realists about others. Structural realism is one such answer and it is best understood and motivated by looking at case study.⁴

FRESNEL AND THE ETHER

The dominant view of the nature of light in the 19th century was corpuscularian. Light consisted of rays formed of tiny particles. This theory could account for a wide range of optical phenomena including, reflection and refraction but it had difficulty in accounting for diffraction effects. Augustin Fresnel, showed that many of these diffraction phenomena (and indeed other features of light like polarization) could be accounted for if we adopted a new theory of light in which it was described as a transverse wave propagating through a luminiferous ether. Henri Poisson an advocate of the rival corpuscularian view, thought this led to a ridiculous result. If Fresnel's theory were right there should be a bright white spot in the centre of the shadow cast by an opaque disc. Poisson took this result to be a reductio of Fresnel's theory. However, an experiment was performed to test Poisson's prediction and surprisingly the white spot was observed. Surely such an incredible result shows that light must be as Fresnel described it. Unfortunately, later physics does not support this view. Subsequent work by Maxwell and Einstein consigned the ether to history. If there's no ether, there's no vibrating in the ether and so there is nothing in the world like Fresnel's description of light.

Structural realists claim we can have the best of both worlds here. If we look more closely at Fresnel's theory we see that certain aspects of it are retained in later physical theories like Maxwell's. Specifically, the mathematical equations

⁴ This is Worrall's (1989) example.

with which Fresnel described the relative intensities of reflected and refracted light reappear in Maxwell's theory unaltered. (See below.) Of course the referents of the key terms are different. For Fresnel the equations described a mechanical oscillation in a jelly-like stuff; for Maxwell and his successors they describe a displacement current in an electric field. But says the structural realist that shows us that what Fresnel got right was the structure of light as encoded in the mathematical equations; what he got wrong was its underlying nature. So we should be realist with respect to the structural claims of science and anti-realist about the claims that go beyond structure.

Fresnel's equations for the relative intensities of reflected and refracted light.

$$R/I = \tan(i-r)/\tan(i+r)$$

$$R'/I' = \sin(i-r)/\sin(i+r)$$

$$X/I = (2\sin r \cdot \cos i) / \sin(i+r) \cos(i-r)$$

$$X'/I' = 2\sin r \cdot \cos i / \sin(i+r)$$

I^2 , R^2 and X^2 represent the intensities of the incident, reflected and refracted beams respectively for the component of light polarised in the plane of incidence. I'^2 , R'^2 , X'^2 the same for the component of polarised light orthogonal to the plane of incidence. The angle of incidence of the beam is represented by i and the angle of refraction by r .

GENERALISING FROM THE CASE

What are we to make of this suggestion about how to respond to revolutionary episodes. One kind of problem which I will discuss in a bit more detail in the next section is something called the Newman problem which threatens the very coherence of structuralism. But before I look at that I want first to consider the general form of the structural realist response to see if even on its own terms it is plausible.

The structural realist recommends belief in the structure of scientific theories because it is claimed these structures are essential in explaining the success of scientific theories and because when we look at the history of science the mathematical structure is preserved (approximately) across theory change. I say approximately, of course, because the Fresnel case is somewhat atypical. It is not generally the case that exactly the same mathematical equations are retained from one theory to the next. But often enough the equations of a new theory can be recovered by some idealization⁵. For example, by setting the speed of light to infinity we can recover the Galilean transformations of Newtonian physics from the Lorentz transformations of special relativity.

So what we should expect when look through the history of science, if the

⁵ One might worry about what counts as an appropriate idealization. Not every way of deriving a set of equations from another could count for that would trivialize the claims of structural realism. But advocates of structural realism have said very little about what should count as an appropriate idealization.

structuralist is right, is that structure is in fact retained from one theory to the next and that retained structure is essential to the success of the past scientific theory. I think this claim is not true and a simple example can show this.

Consider the shift that I alluded to above the move from Newtonian Mechanics to the special theory of relativity. Both of those theories are theories of space and time and as I explained some of the structuralist intuitions do indeed seem to apply to these theories. Consider another important structural continuity between these theories. Both theories describe space (or space-time) as flat. In other words, there is an important structural continuity at the level of global space-time structure. Moreover, it seems essential to both theories success that they represent space-time in this way. It is difficult to make sense, for example, of special relativity's success without reference to Minkowski space-time. Nevertheless(and I hope the lesson is obvious) even though structure here is retained from one theory to the next and even though that structure seems to play an essential role in explaining the success of the theory, we have for familiar reasons good reason to reject realism about the structure. The familiar reasons are of course further reflection on the future development of physics. General relativity supersedes both Newtonian mechanics and special relativity but, of course, it does not represent the structure of space-time as flat.

I suggest a simple example like this should be enough to undermine our confidence in the structuralist strategy. It just does not seem to be true in general that if a structure e is retained from one theory to the next and that structure is part of what explains the success of the theory then that structure picks out something real. The history of science should make us as chary of structural realist claims as any other.

OSR DOES IT EVEN MAKE SENSE?

So far I have argued that the arguments which are meant to motivate structural realism that appeal to metaphysical underdetermination in physics and theory change in science are no good. The first at best motivates an epistemological position rather than a metaphysical position and the second does not accurately reflect certain key developments in the history of physics. I want now to turn to a much more profound problem from structural realism. I will argue that not only are the arguments for structuralism no good but in fact its key claims make no sense.

First let me begin by considering the position discussed above, epistemic structural realism. What exactly does it mean to believe only in the structure of the theory? Well, the clearest way to make sense of the idea of a mathematical structure is in set theoretic terms. A set theoretic structure asserts that there exists a certain relation or relations defined extensionally which satisfy a domain of objects. But this claim cannot be what a scientific realist wants as a description of what is held true in a scientific theory for it is too easy to satisfy such a structure. Any set of objects provided it has the right cardinality can

satisfy such a structure. This point was first made by W.H. Newman against a

similar proposal from Bertrand Russell⁶. Newman puts it like this:

No important information about the aggregate A, except its cardinal number, is contained in the statement that *there exists* a system of relations, with A as a field, whose structure is an assigned one. For given any aggregate A, a system of relations between its members can be found having any assigned structure compatible with the cardinal number A. (Newman 1928, p.140. Italics in the original)

To know then the structure is to know no more the cardinality of your domain. A scientific realism based on this claim is empty.

Many people (including me) think this objection fatally undermines epistemic structural realism but advocates of the position we are interested in discussing here, OSR, have a very neat solution to this problem. Since according to OSR there are no objects, it makes no sense to talk of the same structure satisfying different domains of objects; all there is is the structure itself. This neat solution though quickly gives rise to an equally serious problem for advocates of OSR.

Advocates of OSR like Ladyman and Ross(2007) think the fundamental structures which exist in the world are mathematical in nature That's why they say things like this:

if one were asked to present the ontology of the world according to ... [general relativity] one would present the apparatus of differential geometry and the field equations and then go on to explain the topology and other characteristic of the particular model... of these equations... There is nothing more to be said" (159)

But now we have to answer the obvious question how is this physical structure different from the mathematical structure which we would normally think of as representation of it. What in short makes something a physical structure as opposed to a mathematical structure? The problem for a structuralist metaphysics is that whatever they appeal to differentiate physical from mathematical structure they will face the following dilemma: Let us call the fact which differentiates physical from non-physical structure, X. Is X a structural fact? If not then there is at least one non-structural fact and so OSR is false. If it structural, then we have just pushed our question back. How now are we meant to distinguish this enriched structure from an enriched mathematical structure which could represent it?⁷

Both Ladyman and Ross (Ladyman, Ross and Kincaid 2013, ch.6) in recent talks and papers have tried to address this issue. Ross has claimed the world is the totality of non-redundant statistics and that this avoids the above dilemma since there is "no such thing as purely formal statistics". But this is just rhetoric. Statistics can indeed be worked out purely formally (that is what they study in

⁶ Russell (1927) is taken by advocates of SR as an early version the structuralist project although its motivations are quite different from those of Worrall's paper. See Stathis Psillos (2001) for what he calls the upward and downward paths to SR.

⁷ Disappointingly, despite being aware of the question, Ladyman and Ross (2006) in their book "refuse to answer it" (158). But the question is compulsory. Without an answer to it, it is impossible to make sense of the claims of OSR.

departments of pure statistics) what is needed is some reason to say a particular model is an adequate representation of some facts; and that requires a way of differentiating the statistical model from the represented facts. No appeal to non-redundant statistics does that. This is essentially a repacking of the original problem.

Ladyman (2011) in a discussion with van Fraassen offers an alternative route. He has suggested that we might adopt “intensionalism about the relevant relational structure” or “that there is in the world some causal or nomological structure that is represented by logical and mathematical relationships in our theoretical thought”(421). The suggestions here are very vague and programmatic to say the least but I will content myself with two observations. First, this appears to admit the need for non-structural elements in order for there to be a viable position but secondly and this takes us back to where we began, it seems very likely that in order to make sense of and develop these ideas further Ladyman will have to appeal to a priori metaphysical theorising. Accounts of intensional properties or causal structure will not be found to be read off from our science and so OSRists claims to be doing an entirely novel naturalised form of metaphysical theorising will be undermined. They, like more orthodox metaphysicians, will be left to appeal to intuitions and non-empirical data to support their programme.

CONCLUSION

OSR is a bold and ambitious metaphysical programme which promises to offer a new and interesting way to do metaphysics by engaging both with physical theory and its historical development. Anyone who thinks of themselves as a naturalist will find much to celebrate in the work of Ladyman, French, Ross and other advocates of OSR. Unfortunately, the arguments offered in favour of OSR are disappointing. Naturalised metaphysicians have yet to offer us a world view which can be said to be truly naturalistic while acknowledging our epistemic frailty in the face of scientific revolutions. The search must go on.

THALES OF MILETUS AND THE GREEK NATURALISM

JOÃO HOBUSS

1. This essay is not aimed at dealing with any theoretical aspect or vertical discussion concerning naturalism, regardless of the precise meaning this conception has, since, as recognized by Papineau, the term has no precise meaning in contemporary philosophical debates (2009, p. 1). Therefore, considerations about the separation between an ontological and a methodological component in naturalism will be left aside. Why?

There is a simple reason for that: Naturalism interests us here only in its more generic sense, which may bind it to what we could understand as its inception, that is, the philosophical movement occurred in Miletus in the 6th Century BC that had as its most prominent proposer Thales of Miletus, the forerunner of the natural sciences who apparently, anachronism aside, was generally imbued with some conditions inherent to contemporary naturalism.¹ These general conditions are clear enough:

“These philosophers [naturalists] aimed to ally philosophy more closely with science. They urged that reality is exhausted by nature, containing nothing ‘supernatural’, and that the scientific method should be used to investigate all areas of reality, including the ‘human spirit’” (PAPINEAU, 2009, p. 1)

Thus, our focus is the physical universe, and nothing within nature may evade the processes that inhere in science—understood as the principle of and condition to knowledge—, be it from the standpoint of chemistry, biology, physics, sociology or psychology etc. Also, in the Milesians’s point of view, particularly in Thales’s, who is the focus of our interest here, there is an absolute rejection, as shown by the citation above, of any recurrence to supernatural realities, immaterial beings or religious-like revelations whatsoever since it is centered in the physical reality and the objective laws commanding it, that is, nothing that goes beyond observable causations and events.

¹ The main interest of this chapter is not to provide a detailed discussion of Thales’s theses, but actually to go through the texts found within his doxography, aiming at structuring, from the reports that have survived to our days, a perspective about the philosophical and physical preoccupations of our first philosopher, highlighting what distinguishes him from his successors and, eventually, showing that his seemingly awkward conceptions may have a certain sense from the point of view of explaining nature.

2. In his book *Consilience. The unity of knowledge*, Edward Wilson describes his dream and path regarding the unity of knowledge (WILSON, 1979, p. 1), and the search for a synthesis of natural science, exploring thinkers that aimed to establish a structure of natural history and focusing mainly on the implications that the idea of evolution has to biology, “and for philosophy. And for just about everything” (1979, p. 4), not as an idle idea (1979, p. 5), but as something that has been vindicated and tested by “experience and logics” (id.). According to Wilson, his thoughts, as a biologist, “traveled along a chain of causal events, from mutations that alter genes to evolution that multiplies species,” and so on, experiencing what he has termed “Ionian Enchantment” (1979, p. 4). What does this mean?

“It means a belief in the unity of science – a conviction, far deeper than a mere working proposition, that the world is orderly and can be explained by a small number of natural laws. Its roots go back to Thales of Miletus, in Ionia, in the sixth century B.C. The legendary philosopher was considered by Aristotle two centuries later to be the founder of the physical sciences” (1979, p. 5)

Such “Ionian Enchantment,” that dates back to Thales, but also to Anaximander and Anaximenes, is part of a well-defined philosophical context, namely the School of Miletus, which, as much as contemporary naturalists, sought to explain reality and everything that exists from natural phenomena, not taking into consideration any supernatural basis for their hypotheses and arguments, as questionable and naïve as these hypotheses and arguments may be. In other words, they preferred “a search for objective reality over revelations in another way of satisfying religious hunger” (1979, p. 7).

3. The Milesian School represents a break with the mythical cosmogonies and cosmologies, especially due to its *naturalistic approach* that takes into account the various aspects of the constitution of the world such as, for instance, the generation of living beings and the phenomena of the world, as can be seen in Anaximander (ALGRA, 1999, p. 48), ignoring the anthropomorphic gods as “explanatory factors” (id.), satisfied with a free theoretical activity, elaborating hypotheses that completely diverge from those of his predecessors (ALGRA, 1999, p. 49).

Then, what should be highlighted—and this is the heart of what Wilson calls the “Ionian Enchantment”—is the existence of a systematic, harmonic preoccupation with nature, aiming to understand the “general formation and structure of the world” (DONINI; FERRARI, 2012, p.17), especially regarding the “physical and biological processes” that take place in the world, in an attempt to account for the true nature of things. For such, it was absolutely necessary to suppress, as mentioned above, the mythology of cosmologies and cosmogonies (DONINI; FERRARI, 2012, p.17).²

² These first two paragraphs of item III may be found with further development in HOBUSS, J. *Introdução à História da Filosofia Antiga*. Pelotas: NEPFIL/UFPel, 2014, p. 25-28.

Not taking into consideration the compared quality of the thought of the three Milesian philosophers, or *physics*, and their heritage, the figure that excels them in originality is that of Thales of Miletus, the first philosopher, who inaugurated a new way of approaching reality, not only beyond the mythical narrative, but clearly breaking with it:

“Some think that the ancients who lived long before the present generation, and first framed accounts of the gods, had a similar view of nature; for they made Ocean and Tethys the parents of creation, and described the oath of the gods as being by water, which they themselves call Styx; for what is oldest is most honourable, and the most honourable thing is that by which one swears. It may perhaps be uncertain whether this opinion about nature is primitive and ancient, but Thales at any rate is said to have declared himself thus about the first cause” (*Met.* 983b27-984a3).

Aristotle emphasizes, in his *Metaphysics*, the feature that Thales shares with his Milesian colleagues, that is, that the principles from which all things originate are properly material principles, responsible for generating all there is in reality, all that is:

“Of the first philosophers, most thought the principles which were of the nature of matter were the only principles of all things; that of which all things that are consist, and from which they first come to be, and into which they are finally resolved (the substance remaining, but changing in its modifications), this they say is the element and the principle of things, and therefore they think nothing is either generated or destroyed, since this sort of entity is always conserved, as we say Socrates neither comes to be absolutely when he comes to be beautiful or musical, nor ceases to be when he loses these characteristics, because the substratum, Socrates himself, remains. So they say nothing else comes to be or ceases to be; for there must be some entity—either one or more than one—from which all other things come to be, it being conserved” (983b6-983b18).

Among the forerunners of philosophy, Thales was the first to ask about the principle (*archê*) that originates all things, the nature (*phusis*) that underlies all things. The answer is known by all: the *archê* of the whole of reality is water (*hudôr*). Aristotle attributes to Thales the genesis of this kind of perspective. Thales, “the founder of this school of philosophy [or natural science], says the principle is water (for which reason he declared that the earth rests on water).”³ According yet to Aristotle, the idea that water is the principle of all things is due to the fact that the nourishing fountainhead of things is the *moist*: “the nutriment of all things is moist, and that heat itself is generated from the moist and kept alive by it (and that from which they come to be is a principle of all things).” Evidently, Thales provided reasons for sustaining that the principle of all things was water, reasons that do not seem to us as sufficient, or well-grounded, but that support, from a rational argument, his core point: “He got his notion from

³ Cf., also, Diógenes Laércio (DK A I).

this fact, and from the fact that the seeds of all things have a moist nature, and that water is the origin of the nature of moist things” (983b19-983b28).

Aristotle’s testimony is confirmed by Simplicius (DK4 A XIII) when he affirms that “among those who declare the principle as one (cf. also Aëtius in DK A XIIIb) and moving, whom Aristotle names physicists proper, some he claims to be limited, as [does] Thales.” And this principle is the water due to the fact that empirical observation shows that “heat takes its life from moist, the corpses necrose, dry, the seeds of all beings are moist and all feeding is juicy” (...), since water is the principle of moist nature that comprises in itself all things.”

This new way of considering nature from an eminently rational view that we find firstly in Thales, as mentioned by Jonathan Barnes, is what should be celebrated (1982, p. 4) insofar as “they sought out and drank from the springs of reason” (idem). What would it mean? Not only an antagonism between mythical and philosophical narrative,⁵ since it would be a too restraining conception, but above all a background change, wherein “unargued fables were placed by argued theory, that dogma gave way to reason” (idem), that is, despite a doctrinal inconsistency, what should be preserved is the idea that his doctrines, mainly the one concerning Thales, presupposed a rational argumentation, an explanation; and his conclusions, even if ‘simplistic’ or ‘nonsensical’,⁶ followed from it.⁷ Aristotle himself (HUSSEY, 2009, p. 8) was suspicious about the kind of theory developed by Thales and his colleagues, considering it somewhat strange. Nevertheless, he saw in it “at least in intention, a genuinely ‘scientific’ explanation,” which attempted to rationally account for the whole of the universe, understand its functioning and the laws that command it.

⁴ Diels and Kranz (see Bibliography).

⁵ According to Hussey, “It is clear that this sort of general and abstract claim is not the kind of thing that one meets with, either expressed or latent, in Homer or Hesiod. It is characteristic of a theoretical enterprise. Aristotle elsewhere (e.g. *Met.* B.4, 1009a9-19; *Meteor.* II.1, 353a34-b5) contrasts with scientific explanations the explanations of those he calls “writers about gods” (*theologoi*), those who speaks “in myths” (*muthikôs*), implying that they are incomplete, unsatisfactory and not wasting time on.” (HUSSEY, 2009, p. 8)

⁶ Thus Seneca considers the thesis that the earth rests or floats on water: “Thales’s thesis is absurd; indeed, he affirms that the earth rests on water and floats as a ship and that, when one says that the earth shakes, this is due to the mobility of water. No wonder the liquid element is abundant and gives birth to rivers, since the world in its entirety is found within water” (DK A XV). On the idea that the earth flows on water, see Aristotle and Simplicius (DK A XIV).

⁷ Within this view, there is no room for attributing it to “the will - or the caprice - of the gods” (BARNES, 1987, p. 16). There is no room for external interventions insofar as all the natural phenomena or events should be systematically and methodically explained (idem, p. 17): the “thunder was explained scientifically, in naturalistic terms - it was no longer a noise made by a minatory Zeus” (idem, p. 16).

This is precisely what Thales sought to understand, that is, how “that which exists” works, as well as the laws inherent to this process,⁸ as acknowledged in this testimony by Apuleius (DK A XIX):

“Thales of Miletus, undoubtedly the most eminent of the famous seven sages, was the first among the Greek to discover geometry, to show with regard to nature a *firm scientific curiosity* and to observe the asters with great competence, he did important discoveries⁹ thanks to small segments (*parvis lineis*): the marching of the seasons, the courses of the winds, the movement of the stars, the roaring sound of thunder, the oblique orbit of the stars and also the growth of the Moon when it rises, its decline when it hides and the obstacles responsible for its eclipses.”

This kind of experimental attitude may be perceived in the anecdote told about Thales—when he was ridiculed due to his poverty—and the supposed uselessness of philosophy. Utilizing his knowledge of astronomy, he foresaw a successful harvest of olives, which led him to rent all the presses in Miletus and Chios. His prevision ultimately brought him wealth. The lesson that he wanted to teach is that a philosopher—if he had this purpose—could easily become a wealthy man, but this should not be the aim of “his virtuous efforts” (DK A X). His astronomical knowledge, and the great interest aimed at investigating nature, contributed a supplementary anecdote strictly consonant with his scientific curiosity. Having fallen into a hole for being occupied with “the heavenly things,” he was the object of laughter by a Thracian servant because he was so concerned with knowing what happened in the heavens and did not pay attention to what was right in front of him or at his feet (Plato, *Teeteto* 174a).

The reading of Thales’s doxography may seem strange concerning his affirmations that (i) inanimate beings have souls, using as examples magnet and amber or that (ii) all is filled with gods (or *daimones*¹⁰). Both views bring about the suspicion that Thales would not share the contemporary naturalists’ conception that we should focus only on explanations that do not appeal to the existence of supernatural entities or to immaterial minds, thus evading the study of nature and the laws therein originated, a necessary condition for the naturalistic argumentation. But could we treat both views in this sense or should they be explained without the need of an immaterial mind or a supernatural being?

⁸ “The world [for the pre-Socratics] is orderly without being divinely run. Its order is intrinsic: the internal principles of nature are sufficient to explain its structure and its history” (BARNES, 1987, p. 17). Therefore, the universe is an orderly whole that can be understood (*idem*).

⁹ According to Diogenes Laërtius (DK A I), Thales was regarded as the first philosopher, or physicist, to deal with astronomy. He predicted sun eclipses and solstices. It is actually well-known in his doxography that the prevision of an eclipse in the time of Alyattes II was attributed to Thales—not to Darius, as seen in DK A II. On the eclipse, see also DK A V.

¹⁰ “In Homer and others early authors, gods, even Olympians, could be referred to as *daimones* (in: *The Oxford Classical Dictionary*. 3.ed. Oxford: Oxford University Press, 1996, p. 426. Another possible translation would be ‘spirits’.

There are two specific passages about (i); one by Diogenes Laërtius, the other by Aristotle:

“Aristotle and Hippias say that [Thales] gave inanimate things too a share in soul (*psuchê*), taking his evidence from the magnetic stone and from amber” (DK A I).

“It seems, from what they report, that Thales too supposed the *psuchê* to be a sort of motor, given that he said that the magnet has a *psuchê* because it moves iron” (DK A XXII).

Johnathan Barnes always reminds us that, though simple, the arguments are always rational (1982, p. 5) and, in this particular case, they reveal “a keen philosophical eye,” whose internal consistency lends some plausibility to the propositions, exposing an internal consistency (p. 7). The argument could be synthesized in three parts:

- (1) If anything has a motor, it has a *psuchê*. According to Barnes, these are self-starting motors, which Aristotle names “*animators* or *psuchai*” (p. 6 - 7);
- (2) Magnets and pieces of amber have motors. This premise would be, still according to Barnes, a common-sense observation: “magnets or pieces of amber are seen to possess the power to cause locomotion in other things and to move themselves” (p. 6 – 7);
- (3) Magnets and pieces of amber have a *psuchê*. This conclusion follows from (1) and (2): “magnets and pieces of amber are animate things (...) they are alive” (idem).

From this perspective, the building of the argument is not bizarre, or an unpardonable theoretical insult, but it follows a reasoning that presents an internal plausibility that should not be simply discarded as nonsense.

In relation to (ii), which says that everything is filled with gods (cf. DK A I [27], A III, A XXII, A XXIII), there is also a possible, if extravagant, rational and credible explanation that answers well to the reductionist perspective that sees these fragments as an appeal to a supernatural, divine entity. What would this explanation be?

This proposition could be associated to what was just talked about, that is, that, in affirming that all the things in the universe are filled with gods, *panta plêrê theôn einai* (DK A XXII), or *daimones* (DK A I 27 [*kai tôn kosmon empsuchon kai daimonôn plêrê*], A III [*ton de kosmon empsuchon ephê kai daimonôn plêrê*] and A XXIII, which comprehends both terms—gods and *daimones*, when God is mentioned as the world’s mind [*noun tou kosmou ton theon, to de pan empsuchon hama kai daimonôn plêres*]: the two forms appear in the fragments found within DK. Thales simply affirms that all things have a soul (a self-starting motor), that is, a life principle or, as mentioned by Barnes, an animator, which indicates that all things are endowed with soul in an entirely different way as posited by the mythic tradition. Such an idea can be found in Aristotle’s *De Anima* 405a19-21 (DK A XII):

“From what people say about him, it seems that Thales supposed that soul is some kind of moving principle – if, that is, he said that the magnet has a soul because it moves iron.”

As said by Algra (1999. P. 53), Thales is positing that the physical universe in its entirety has “a principle of movement, even in inanimate objects,” and that this principle of movement may be named *psuchê* (or even God or gods). Algra concludes that there would be some conception of divinity left in his cosmology, but with a very important remark:

“Even if this shows that the world picture of the early Milesians was not fully ‘secularized’, it should be stressed that instead of the more or less anthropomorphically conceived cosmic deities of Hesiod we now have a more despersonalized or ‘physicalized’ conception of divinity that does not readily allow for a description in wholly theistic terms” (ALGRA, 1999, p. 53).

It is my belief that Algra does not take the definitive step in acknowledging the background distinction between the mythical and the philosophical views, since he does not break with the existence of a certain notion of the divine, even if it was not fully secularized’ or ‘a more despersonalized or “physicalized’ conception of divinity’ and ‘not a description in wholly theistic terms’. The use of god or spirits seems to have a merely allegorical meaning that does not authorize even a distant relationship with the mythical narrative. In reality, the breakage with mythical narrative is clear and even with the instrumental use of terms belonging to this narrative, the perspective is absolutely different.

Does this allow treating Thales as the earliest naturalist philosopher? Maybe it is a bit of an audacity and an anachronism, but not a crime, regarding him, at least, as a proto-naturalist in his purposes, since his investigations aimed simply at explaining nature focusing only on rationally-grounded explanations, attempting to understand the way the natural laws and events functioned. Due to this, the various reports consider him the founder of the studies on nature, which is the special niche occupied by him in the History of Philosophy:

“Thales is reported to be the first to reveal natural history to the Greeks, even though many before him were focused on it, as it seems to be the case also of Theophrastus; but he is in such a manner above them that he has eclipsed all those who preceded him” (Simplicius *in* DK B1).

THE NATURAL KIND STATUS OF DELUSION

JOSÉ EDUARDO PORCHER

INTRODUCTION

Delusion is defined by the most recent edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) as a ‘false belief based on incorrect inference about external reality that is firmly held despite what almost everyone else believes and despite what constitutes incontrovertible and obvious proof or evidence to the contrary’ (American Psychiatric Association 2013, p. 819). Predictably, a great variety of phenomena are apt to be grouped under such a definition. Indeed, people who are deemed to be clinically delusional affirm many different things in many different contexts. Here are some of them (Davies and Coltheart 2000, p. 1):

‘My closest relatives have been replaced by impostors.’

‘I am dead.’

‘I am being followed around by people who are known to me but who are unrecognizable because they are in disguise.’

‘The person in the mirror is not really me.’

‘A person I knew who died is nevertheless in the hospital ward today.’

‘This arm [the speaker’s left arm] is not mine, it is yours; you have three arms.’

‘Someone else is able to control my thoughts.’

‘Someone else’s thoughts are being inserted into my mind.’¹

What follows is an investigation about our warrant for grouping such disparate phenomena together. My primary aim is to assess the prospects for a scientific theory of delusion through the examination of the scientific respectability of this psychiatric category—a status which is arguably put in jeopardy by the fact that the detection and attribution of delusion seem to stem not from causal classification but from the application of what we may call ‘folk psychiatry’. I will do so by first introducing the philosophical notion of *natural kind* and examining the question of whether psychiatric kinds as a whole meet the demands required for a kind to be an objective, mind-independent distinction in nature. I will then introduce a liberal sense in which biological taxa as well as

¹ These examples pertain to eight different subtypes of clinical delusion, respectively: Capgras delusion, Cotard delusion, Frégoli delusion, mirrored-self misidentification, reduplicative paramnesia, somatoparaphrenia, thought control, and thought insertion. See Radden (2011) and Porcher (2016) for a more in-depth introduction to delusion.

psychiatric categories might be viewed as natural kinds—namely, the *homeostatic property cluster* model. Subsequently, I will introduce and assess how models of the detection and attribution of mental disorder may impact even a liberal understanding of delusion as a natural kind. Finally, I will conclude by making a case for a folk-psychological understanding of ‘delusion’ in general while also recommending a natural-kind methodology for the investigation of subtypes of delusion.

1. KINDS OF KINDS

Are mental disorders real? One of the main theoretical challenges for psychiatry is to determine whether the *kinds* it investigates are *natural*. Psychiatry’s scientific credentials came under heavy criticism in the 1960’s and 1970’s—the most radical embodiment of which was represented by the so-called anti-psychiatry movement, which questioned whether mental disorder represents the pathologizing of normal problems of living. Thomas Szasz, the father of anti-psychiatry, argued not only that mental disorder as a kind fails to pick a real distinction in nature, but that it is just a ‘convenient myth’ (1961, p. 113). This intuition is reinforced by controversies such as that over the recent removal of the “bereavement exclusion” in the diagnosis of depression in the DSM-5. Likewise the proposed addition of ‘persistent complex bereavement disorder’ in an attempt to classify those who are significantly impaired by prolonged grief symptoms for at least one month after six months of bereavement.² Against the backdrop of challenges to the validity of psychiatric classifications as a whole, the task is to make clear the basis on which conditions are included or excluded from the manuals and why this basis is scientific and objective and not just a matter of social rules of normal behavior (Bolton 2008, p. 164). If entities classified as mental disorders could be shown to be natural kinds, then many of the controversies surrounding the status of psychiatry as a serious scientific endeavor could be resolved. However, this will depend on what exactly one takes natural kinds to be.

1.1 ESSENTIALISM ABOUT NATURAL KINDS

What are natural kinds? What characteristics must a kind have in order for it to be considered a natural kind? The traditional account of natural kinds is represented by various forms of *essentialism* which date back to the Aristotelian tradition, in which essences had both causal and classificatory (sortal) roles. The causal role referred to the underlying properties that determined and sustained an instance’s visible properties. Because these underlying properties were supposed to be fixed, they were identified with the nature of a kind—that which makes it be what it is. After the rise of natural philosophy in the seventeenth century, the essential hidden properties which Locke called ‘real essences’ came to be

² Persistent complex bereavement disorder was placed in the chapter ‘Conditions for Further Study’ in the DSM-5 after its proposed addition generated a great deal of controversy.

identified with underlying structural properties which, he argued, are not observable.³ In the twentieth century, essentialism was mostly related with the revival of the notion of natural kinds in the work of Saul Kripke (1972) and Hilary Putnam (1975),⁴ which followed the skepticism about the stability of scientific knowledge brought about by the work of Thomas Kuhn (1962).

As Marc Ereshefsky (2009) observes, essentialism usually involves three main tenets: first, all and only the members of a kind share a common essence; second, that essence is a property, or a set of properties, that all the members of a kind must have; and third, a kind's essence causes the other properties associated with that kind. So, for example, the essence of gold is gold's atomic structure, and that atomic structure occurs in all and only pieces of gold. That structure is a property that all gold must have as opposed to such accidental properties as being valuable to humans. And the atomic structure of gold causes pieces of gold to have the properties associated with that kind, such as readily dissolving in mercury at room temperature, conducting heat and electricity, and being unaffected by air and moisture.

The reason why it matters for the development of a science that its kinds be natural in the sense of picking up essential distinctions has to do with the fact that such kinds will be ideally suited to figure in key scientific practices such as induction, explanation, classification, and discovery. Natural kinds pick out classes about which non-accidental, scientifically relevant, inductive generalizations can be formulated, since its members share many non-accidentally related properties. The reliably co-varying clustering of properties that instances of natural kinds possess is, however, contingent (as opposed to logically or conceptually necessary) and its existence calls out for explanation, usually undertaken through the identification and specification of the structures, processes, and mechanisms that causally explain the property clusters associated with the kind under consideration.

In other words, one's ability to make inferences about members of a natural kind is explained with reference to their shared underlying properties. Being some such natural kind explains why an instance of that kind has the features that it does, and that explanation is to be found in studying the intrinsic underlying properties an instance shares with other instances of that kind. Furthermore, with respect to the classificatory role, if one can identify the essence of a thing, one may be able to determine its place in the natural order. According to essentialism, if you want to know whether something is a true member of a natural kind, you should check whether the causally essential underlying properties are present, as such properties will invariably be necessary and sufficient conditions for membership in a natural kind. Thus, essentialism implies that there is a correct classification of naturally occurring kinds out there

³ It is fair to say that Locke underestimated the kinds of observation that technology would eventually allow us to make of properties which are potentially essential, such as the number of protons in the nucleus of an atom, or the genetic code in specific DNA sequences.

⁴ But see Hacking (2007) for criticism of the lumping together of Putnam's and Kripke's theories of natural kinds on the basis that Putnam was not an austere essentialist.

waiting to be discovered. As the philosophical adage goes, nature is such that it can be “carved at its joints.”

Besides figuring in the practices of generalization, explanation, classification, and discovery, Richard Samuels (2009) points out three further characteristics that flow from natural-kindhood as necessary conditions for the scientific respectability of any given kind. Given that natural kinds possess a *sortal* essence,⁵ they will be *discrete* classes of entities that can be clearly demarcated from other phenomena and they will be highly *homogeneous* classes as well. Moreover, natural kinds will be *mind-independent* in an important sense,⁶ which Sam Page (2006) calls *individuating independence*, namely, that of being circumscribed by boundaries that are totally independent of how we categorize things. Page illustrates his concept by alluding to the individuation of the night sky into constellations: ‘Though it is *prima facie* plausible that reality is individuated intrinsically into stars, reality is not individuated intrinsically into constellations, since it is people who divide the night sky into constellations’ (2006, p. 328).

Essentialism about psychiatric kinds—the view that psychiatric disorders are (or at any rate should be) akin to stars, not to constellations—is associated with the biomedical model of psychiatry, which proposes that psychiatric kinds can and should be isolated by studying underlying biopathological processes. Jerome Wakefield’s (1992) *harmful dysfunction* model, arguably the most important philosophical theory about the nature of mental disorder, recognizes the claims of Szasz and others concerning the evaluative nature of psychiatric diagnosis without thereby abandoning realism about psychiatric disorders. Wakefield argues that the presence or absence of a dysfunction is a factual matter, just as the presence or absence of a natural function is. Since natural functions were selected for during evolution because of their contribution to the survival of the organism, evaluative statements about functions (and, hence, dysfunctions) can be translated into objective, factual statements about evolutionary history. To qualify as a “disorder,” however, Wakefield acknowledges that there must also

⁵ As Samuels (2009, p. 57) uses the term, sortal essences consist of intrinsic properties and, as a matter of metaphysical necessity, they are possessed by all and only the members of the kind. *Causal* essences, on the other hand, do not imply these commitments, and are simply the set of properties that figure in causal explanations of a given kind. So all sortal essences are causal essences but not vice versa.

⁶ Following Page (2006), Samuels (2009, pp. 53–4) identifies three possible senses of mind-independence that do *not* flow from natural-kindhood and are, therefore, irrelevant to the characterization of natural kinds. The first is that attached to theoretical entities (e.g. quarks, electrical fields, and chemical compounds), which should not be considered trivially mind-independent, non-natural kinds. The second is that attached to entities whose existence metaphysically necessitates the existence of minds, such as psychological kinds as beliefs, desires, delusions, etc. and, again, should not be considered trivially non-natural. Finally, and perhaps more controversially, Samuels rejects the relevance of causal dependence on mental activity, which is true of such kinds as toy poodles and the radioactive chemical element californium, as he argues that this feature should not trivially imply that such kinds are not “natural” in the *scientifically* relevant sense (i.e. though not naturally-occurring, they may nevertheless turn out to figure in all relevant scientific practices).

be evidence that the condition in question is *harmful* to its bearer—and this will be an inherently evaluative, normatively assessable aspect of all judgments of pathology.

Given the present stage of development of biological psychiatry, however, the essences of the dysfunctions that constitute psychiatric disorders—alongside the evaluative aspect of suffering or impairment—are yet to be discovered, just as the essence of electrons and gold once were. Until the necessary scientific discoveries are made, their essences are, so to speak, in a black box. As Peter Zachar explains, Wakefield's (2004) black-box essentialism follows the scenario proposed by Putnam and Kripke wherein, at some point in history, there occurs a "baptismal" event in which, in the example at hand, a disorder is clinically observed and named: "This is psychopathy," said Hervey Cleckley (1941). 'This is autism,' said Leo Kanner (1935). If the original disorder concept can be developed into a proper scientific construct (one based on an objective dysfunction), the clinician's original concept can be said to have indirectly referred to the objective dysfunction all along' (2014b, pp. 83–4).

Note, however, with respect to the aforementioned conditions for the scientific respectability of a kind, that biological taxa such as species appear to meet all of them and, still, they are widely regarded as failing to constitute essentialistic natural kinds⁷ as do chemical kinds such as ascorbic acid and H₂O, and physical kinds such as quark and lenticular galaxy. This is the case because, as the first tenet of essentialism requires, for a biological trait to be the essence of a species that trait must occur in *all and only* the members of that species. However, as Ereshefsky (2001, p. 98) points out, a number of biological forces work against the uniqueness and universality of a trait in any given species. For example, suppose a genetically-based trait were found in all the members of a species, such as the unique genetic code of lemons that Putnam (1975) speculates is the essence of lemons. The forces of non-adaptive causes of evolution such as mutation and genetic drift can cause the disappearance of that trait in a future member of the species. Furthermore, as Ereshefsky observes, even if a trait occurred in all the members of a species, that trait would be the *essence* of a species only if it were unique to that species. But organisms of different species often have common traits because they inherit similar genes and developmental resources from common ancestors. Therefore, given the requirements of essentialism and the forces of evolution, essentialism about biological kinds has been widely rejected.⁸

⁷ From now on, I drop 'essentialistic' as always refer to natural kinds in the essentialistic sense unless otherwise noted. As we will see below, the term 'natural kind' has been re-appropriated by authors who believe that essentialism is too stringent, while believing that less stringent criteria can properly characterize kinds as 'natural' (Boyd 1991).

⁸ Three main views have been advanced in response to this: denying that species are natural kinds and looking elsewhere in biology for kinds with essences (Hull 1978); arguing that species are indeed kinds with essences, but that their essences are of a non-traditional variety (Okasha 2002); and, as we will see below, arguing that natural kinds do not require the sort of essences implied by essentialism (Boyd 1999).

If biological kinds are not amenable to conceptualization as natural kinds, then what chance do psychiatric kinds stand of successfully being characterized as such? Zachar (2000) argues that conceptualizing psychiatric disorders as bounded entities in nature is inconsistent with evolutionary biology's understanding of species. Indeed, as Nick Haslam (2014, p. 11) notes, psychiatric classification would be a great deal easier if its diagnostic entities were like biological species, since, while the process of demarcating biological taxa rests on the scientifically impeccable confidence that naturally occurring biological kinds exist, the taxonomic situation in psychiatry is very different, as mental disorders do not pick out distinct, reproductively isolated, spatially concentrated populations. Moreover, while biological species are "indifferent kinds", at least some mental disorders seem to be "interactive kinds" (Hacking 1999), since those who are classified are often aware of being labeled and may come to change their behavior and even their self-experience in consequence of such awareness, thus producing a "looping effect" whereby the labels may change in virtue of their subjects changing (Hacking 2007b).

Furthermore, in stark contrast to their biological counterparts, psychiatric kinds (and *kinds of people* more generally) tend to be at least partly shaped by social processes and normative concerns. These considerations are the motivating force behind the anti-essentialist argument in philosophy of psychiatry. As we will see, the cogency of this argument will depend on how exactly one should understand 'essence', as essentialism about natural kinds has been challenged in recent years (Boyd 1991). Also, it will depend on the plausibility of the repudiation of pluralism—the view that different psychiatric kinds differ in how much they fail to meet the criteria for natural-kindhood (Haslam 2002)—the acceptance of which would in principle keep open the possibility that at least *some* mental disorders might have essences. For now, however, I will assume that the general argument is cogent in order to consider what may be proposed instead to properly capture the features of psychiatric kinds, noting that by assuming that they are not natural kinds one is not immediately committed to the view that they are non-kinds (*pace* Szasz). Following a nuanced classification of kinds of kinds, such as that offered by Haslam (2014), will go a long way toward disabusing one of the notion that distinctions proper must be essential or fail to be real distinctions at all. His schematic account is based on five kinds of kinds that satisfy increasingly stringent criteria, each successive kind of kind having to meet one more requirement, with natural kinds being on the top of the ladder.

In the remainder of this section, I will go over the different kinds of kinds that fall short of being distinguishable by a category essence: dimensions, practical kinds, fuzzy kinds, and discrete kinds. I will connect these notions to the discussion of natural-kindhood in the philosophy of psychiatry, as well as to the more general discussion of the proper way to characterize natural kinds, within which the most widely adopted view states that natural kinds should not be conceptualized essentialistically, but in terms of property clusters sustained by complex, mutually reinforcing networks of causal mechanisms.

1.2 DIMENSIONS AND PRACTICAL KINDS

The first kind of kind and the least demanding structure in Haslam's model is what he refers to as *dimensions* (strictly speaking a non-kind, since they do not define delimited categories). The label comes from the standard categorical/dimensional distinction in psychopathology research and theory, motivated by the categories of personality disorder which, perhaps more than any other current DSM category, do not seem to be distinct species (Clark, Watson, and Reynolds 1995; Livesley 2003; Widiger and Sanderson 1995). Zachar (2014, p. 93) alludes to a model introduced by Livesley (2003), in which once the pathological dimensions have been identified—which may include narcissism, impulsivity, anxiousness, social detachment, and hostility (Widiger, Livesley, and Clark 2009)—patients meeting criteria for a broad category called 'personality disorder' are distinguished from one another by their respective position on the dimensions. To qualify as a dimension, all that is required for a kind, such as any given mental disorder, is that there be a set of correlated properties, such as symptoms. As Haslam puts it, 'Individuals may differ by degree along a dimension by possessing greater or lesser numbers or degrees of these properties. Variation along a dimension is continuous and seamless, so there is no naturally occurring break separating individuals who are affected with a condition from those who are not' (2014, p. 14). In other words, if psychiatric kinds were dimensions, this would amount to there not being delimited conditions at all. A cutpoint would be defined on the dimension so that the quantitative variation would be simplified into a dichotomous diagnosis, but its placement would be arbitrary.

Thus, proponents of dimensional models of psychopathology hold that the distribution of variation on psychopathology-related dimensions is continuous in the same sense as what philosophers refer to as 'vague predicates'. These models are devised in response to the limitations of the purely categorical approach, such as the failure to capture individual differences in disorder severity, and clinically significant features subsumed by other disorders or falling below conventional DSM thresholds (Brown and Barlow 2005). Nevertheless, while rejecting the view that psychiatric kinds are natural kinds, Zachar (2000) argues that mental disorders pick out reasonably stable, nonarbitrary patterns that can be identified with varying levels of reliability and validity, and that the application of many of the distinctions of psychopathology is justified by its usefulness for clinical purposes, being demarcated on the basis of external considerations rather than on the basis of internal discontinuities. In keeping with these observations, Zachar proposes that mental disorders be conceptualized as *practical kinds*, the next rung in Haslam's ladder, which refers to the least demanding sort of non-arbitrary cutpoint—that of pragmatically grounded distinctions.⁹

As an example from outside the field of psychiatry, Zachar (2014b, pp. 154–5) alludes to the distinction between an adult and a child. Although the kinds

⁹ Though, as we will see below, Zachar's most recent proposal acknowledges the middle way between practical kinds and essentialism about natural kinds embodied in Richard Boyd's property-cluster approach, going so far as to state that Boyd's model is probably the most appropriate for conceptualizing most psychiatric disorders (Zachar 2014, p. 94).

‘adult’ and ‘child’ are not in themselves sharply demarcated, the uses for which we deploy them will determine where their boundaries should be drawn. Consequently, many distinctions between adults and children are context-dependent. For example, if our aim is to decide who is able to vote, engage in consensual sex, get married, be sent to prison, drink alcohol, or enter into a legal contract, each of those considerations will result in different ways of demarcating adulthood (Horwitz and Wakefield 2012, p. 53).

As medical examples of non-arbitrary cutpoints on continuous dimensions, Haslam (2014, p. 14) points out blood pressure values for diagnosing hypertension and Body Mass Index values for diagnosing obesity—values that roughly correspond to levels at which health risks become more likely. When at some point along a dimension the severity of the relevant symptoms becomes clinically significant or a source of functional impairment, the existence of a non-arbitrary, pragmatic distinction is justified.

So practical kinds, while fuzzier than natural kinds, are not open to the charge of arbitrariness as dimensions are (at least as conceptualized in Haslam’s model). The classification of practical kinds requires balancing criteria that do change their values in different contexts depending on treatment goals, research priorities, and disciplinary standards of validity. As a consequence, practical kinds fall short of possessing the perfect reliability one may be justified to expect from natural kinds. Relating the practical-kinds model to his claim that psychiatric nosology is inherently goal-oriented, Zachar has recently elaborated on the dynamics of classification within his model, observing that it emphasizes that discovery of fact contributes greatly to progress in classification, but that discovery alone cannot tell us how to classify: ‘For example, discovering that a mild form of cognitive disorganization (schizotypy) is common in families of people with schizophrenia was an important finding that highlighted an objective feature of the world. Should schizotypy, therefore, be classified as mild manifestation of a unitary schizophrenic spectrum (a genetic grouping)? Another possibility is that should it be classified as a premorbid personality style that represents a vulnerability to the mental illness of schizophrenia. In which box should it be placed?’ (2014, p. 90). Zachar’s point is that, apart from goals relating to classification and theory-building, neither demarcation is privileged in and of itself.

The presence of goal-oriented cutpoints raises the question of whether practical kinds are apt to count as scientifically relevant kinds, and this, in turn, raises the question of the minimal criteria of scientifically-relevant kindhood. Zachar defers to Nelson Goodman, who did not advocate for natural kinds or scientific realism, but instead offered a theory of relevant kinds. With respect to the criteria for relevance, according to Goodman, good scientific kinds *support induction* (to a greater or lesser degree) or, as he would later put it, they have properties that are “projectible,” meaning that if we observe certain properties in a subset of a kind, we can infer that these properties will occur in other instances of the same kind, allowing us to confirm generalizations about that kind (Goodman 1978, 1983). Let us assume, for the sake of the argument, that projectibility is a good enough criterion of relevance. Do psychiatric kinds support induction? Even though present classifications of mental disorders are

highly variable with respect to validity, and in spite of diagnosis being presently based on polythetic categories,¹⁰ research on mental disorder has been able to produce many useful generalizations.¹¹ The question is whether these generalizations are based on (at least some) psychiatric kinds being held together by shared *causal* mechanisms or if they are based solely on these kinds's shared surface features, meaning that they are merely practical kinds.

The practical-kinds model is implicit in the symptom-based nosologies of current diagnostic manuals which aim at grouping patients into useful classes that serve practical goals (such as predicting behavior, assessing genetic risk, or selecting a course of treatment). This grouping, effective as it may be, does not require that diagnoses be grounded in shared causal processes. On the other hand, the assumed causal heterogeneity of psychiatric kinds does not immediately imply that they cannot be causally classified. Note, however, that as the existence of shared causal mechanisms underlying mental disorders is currently an open question, assuming that a causal classification of psychiatric kinds is tenable is something of a “black box” approach (as is Wakefield's harmful dysfunction model). Nevertheless, as Kenneth Kendler, Peter Zachar, and Carl Craver (2011) argue, by focusing solely on the adjustments and compromises that actually occur in classification, the practical-kinds model fails to suggest a way toward progress. In other words, the model is purely descriptive of the current state of psychiatric classifications. If progress is to be made, however, linking disorders to their etiology and underlying mechanisms is indubitably psychiatry's best bet. For this reason, psychiatry may profit from conceptualizing its kinds in a way that goes beyond the merely pragmatic and

¹⁰ Polythetic (as opposed to monothetic) categories were introduced in the DSM-III (1987) and are still used in the present edition, DSM-5 (2013). Polythetic classification is carried out by assigning a certain number of criteria, of which some, but not all, need to be met in order for an individual to be a member. So, for example, the diagnosis of schizophrenia is partially dependent on the patient showing two or more of the following symptoms (for much of the time during a one-month period): delusions, hallucinations, disorganized speech, grossly disorganized or catatonic behavior, and negative symptoms such as blunted affect, *alogia*, and *avolition*.

¹¹ For example, with respect to depression, preventive efforts result in a decrease in rates of the condition of between 22 and 38% (Cuijpers et al. 2008), and stepped-care intervention (watchful waiting, cognitive behavioral therapy, and medication in some cases) has achieved a 50% lower incidence rate in a patient group aged 75 or older (van't Veer-Tazelaar et al. 2009). With respect to schizophrenia, a combination of new medications and community-case management—a multidisciplinary team of mental health professionals who engage with the patient and their carers inside and outside the hospital, and ensure a combination of health and social care—has resulted in remission of about 80% of patients, especially if treatment is initiated early during the first episode of the illness (van Os and Kapur 2009). With respect to bipolar disorders, prodromal symptoms (i.e. those preceding a relapse) can be reliably identified by at least 80% of individuals with bipolar disorder (Jackson, Cavanagh, and Scott 2003), and teaching patients coping strategies to employ when noticing the symptoms, such as stimulation reduction and seeking professional help, has been correlated significantly with better social functioning (Lam and Wong 2005).

assumes internal (but not necessarily external) discontinuities. To this end, we may climb one more rung in Haslam’s ladder, toward a more ambitious model.

1.3 FUZZY KINDS AND DISCRETE KINDS

Dimensions and practical kinds both represent forms of continuous variation. According to Haslam, such variation becomes categorical in a deeper sense when there exists some sort of internal discontinuity within a kind which cannot be accounted for by pragmatic considerations alone: ‘Such a discontinuity involves a break on the underlying continuum, which produces a qualitative distinction between people who fall above the discontinuity and those who fall below it. An example is a threshold effect, in which a qualitative change of state occurs at a certain point on an underlying continuum (e.g., a liquid turning to a gas at a certain temperature, or a spring losing its tension beyond its elastic limit)’ (2014, p. 15). When internal discontinuities within a kind are present but are not sharp, we have what Haslam calls *fuzzy kinds*. Within these, then, kind membership will not always be definite: there will be a penumbra of intermediate cases between those that are definitely members of the kind and those that are definitely not.

On the other hand, when internal discontinuities are sharp but no set of essential properties exists, we step up Haslam’s ladder once again to find what he calls *discrete kinds*. In this kind of kind we have what may properly be called a category boundary. However, Haslam points out that discrete kinds may have a variety of possible causal underpinnings, as many types of causal explanation can yield category boundaries: ‘These causal explanation types include sharp threshold effects (where the qualitative change of state is abrupt), dynamic interactions of multiple causal factors, and explanations that invoke centripetal tendencies within categories (e.g., conscious identification with a group or label) and/or differentiating tendencies’ (2014, p. 15). This immediately makes discrete kinds excellent candidates for scientific respectability in the eyes of those who argue that scientific practice does not require an essence in the traditional sense of a microstructural property that explains all the other properties of a kind while also being unique to that kind.

Indeed, both fuzzy and discrete kinds are candidates for natural kindhood if one refuses to accept that what makes a kind a natural kind is its possession of an essence, rather than its utility in induction and other scientific practices. Within the non-essentialist *kinds-in-science* tradition (Cooper 2013), fuzzy, discrete, and essentialistic natural kinds are all proper subsets of inductively useful kinds.¹² Within this tradition, several accounts of kinds have been developed with the aim of explaining how it is that kinds like biological species—in which there simply are no essential properties to be found—can

¹² Though the history of natural kind thought is usually traced back to Locke’s real essences (Boyd 1991), Murphy (2006, p. 335, fn. 6) notes that, as a historical precedent for the kinds-in-science tradition, Hacking (1991) argues that the notion of natural kinds indubitably surfaces in Mill and Venn in the mid-nineteenth century in connection with induction—something which did not preoccupy philosophers before Hume.

successfully ground explanations and inductive inferences. Insofar as the most ambitious sense in which psychiatric kinds might turn out to be natural is the same in which biological kinds are taken to be natural, such accounts of kindhood are of particular interest for the conceptualization of mental disorders as something belonging between practical kinds and kinds with essences.

John Dupré (1981, 1993) argues for *promiscuous realism*—the view that there are countless, yet legitimate ways of dividing up the world into kinds. He asks us to consider the entities of some domain mapped into a multidimensional space wherein the different dimensions map onto different properties, as in cluster analysis—a statistical method for grouping sets of objects based on their similarities, in such a way that objects in the same cluster are more similar to each other than to those in other clusters. According to Dupré, biological species—as well as higher taxa such as families and kingdoms, and lower ranks such as subspecies and varieties—would be identified with some such clusters. His realism has to do with the fact that he accepts that the world possesses individuals which are objectively similar to each other, sharing properties and, thus, being identifiable as being of the same kind. The promiscuity of Dupré’s realism, on the other hand, has to do with the fact that he denies that these properties are intrinsic properties of kinds and, in line with Haslam’s concept of fuzzy kinds, he argues that natural kinds are not necessarily categorically distinct (i.e., they are not necessarily discrete kinds). Moreover, such taxonomic promiscuity is reflected on our classificatory practices both in the context of common sense and within science.

In the context of common sense, a (presumed) natural kind such as lilies is classified as a flower, although, in biology, species which are commonly referred to as lilies occur in numerous genera of the lily family (Liliaceae), including bulbs such as garlic and onions. However, as Dupré observes, to include the onions and garlincs in the reference of the English word ‘lily’ would surely amount to a debasement of the term (1981, p. 74). The moral is that common sense and biology provide us with pluralistic ways of classifying lilies and each is equally legitimate depending on our interests. This is not to say that Dupré’s kinds are merely practical—it means that his conception of natural kinds takes seriously the different classifications that arise from a variety of interests. Indeed, cross-classification sometimes occurs within the context of a single science, to which the countless ways of classifying species bear witness (Dupré 1993, p. 38).

By denying that there is one unique way of demarcating the set of natural kinds, Richard Boyd (1991, 1999) endorses promiscuous realism. Furthermore, by emphasizing that members of a kind share properties for a reason, his *homeostatic property cluster* (HPC) account elaborates on Dupré’s idea. In a near-consensus in recent philosophy of science, the HPC account has been widely seen not only as the most successful approach to make sense of the intuitive natural-kindhood of biological species, but as quite simply the best account of natural-kindhood (Samuels and Ferreira 2010). The HPC model defers to the kinds-in-science tradition by stating that natural kinds are scientifically relevant kinds and that these are, at a minimum, fuzzy sets defined

by homeostatic¹³ mechanisms at multiple levels that act and interact to produce the key properties associated with the kind. These mechanisms are the reason why members of a kind are, and continue to be, alike. Importantly, they are also the reason why the clusters of phenomena identifiable as being of the same kind are similar enough to be subject to explanation in terms of the same underlying causal properties. Thus, Dominic Murphy (2006, p. 338) refers to Boyd’s account as a refined form of essentialism, since homeostatic properties substitute and play the same role of what in “simple” essentialism constituted the essence of a kind (namely, microstructural properties). By not insisting on necessary properties or a single, essential cause, and by not specifying that such a cause must be biological, the HPC account is clearly broader than simple essentialism and advances a much more liberal sense of natural-kindhood.

So Boyd’s natural kinds are, minimally, fuzzy kinds. In cluster-analytic terms, if the members of different fuzzy kinds whose members share a certain number of properties are plotted in a multidimensional space, there will not always be a clear gap between them. As Haslam (2014, p. 18) notes, since homeostatic mechanisms merely produce correlations among properties and resemblance among entities that possess those properties, there is no reason to assume that similarity-generating mechanisms will always yield sharp discontinuities between entities that possess sufficient levels or numbers of those properties and entities that do not. This leads Carl Craver (2009) to conclude that HPC kinds have a *prototype* or *family-resemblance* structure. Note, however, that both discrete and essentialistic kinds are also proper subsets of the set of HPC kinds, so that Boyd’s account accommodates the intuitively plausible possibility that there are different levels of natural-kindhood—in Haslam’s five-tier classification, these levels comprise all kinds for which there are internal discontinuities independent of our interests. In this way, some scientifically relevant kinds may turn out to be fuzzy, others discrete, and still others may turn out to have essences. For example, membership in the kinds encompassed by chemical elements may be essentially defined by the number of protons found in the nucleus of an atom. This is part of the appeal of Boyd’s account, since there is no reason to think that psychiatric disorders, biological species, and chemical elements must pertain to the same kind of kind, and, according to the kinds-in-science tradition, there is also no reason to deny natural-kind status to non-essentialistic kinds as a matter of principle.

The inductive potential of HPC kinds is underwritten by the fact that if properties are held together homeostatically, then we will be able to conclude on the basis of one property that others will typically occur with it. Boyd’s focus on the underlying causal mechanisms that make homeostasis possible is important for the present investigation because it ties the HPC model to causal explanation and classification which, as we have seen, is absent from the practical-kinds model—the main competing model of psychiatric kinds. As Samuels notes, for any homeostatic property cluster ‘there is some set of empirically discoverable

¹³ Homeostasis being the property of a system or mechanism by which variables are regulated so that internal conditions remain stable and relatively constant.

causal mechanisms, processes, structures, and constraints—a *causal essence*, if you will—that causally explains the co-variation of these various symptoms’ (2009, p. 55). Therefore, kind-membership will be defined not by sets of co-occurring properties or symptoms, as mental disorders are presently demarcated in diagnostic manuals such as DSM-5, but by the set of causal mechanisms that make these properties occur together. On the other hand, psychiatric conditions could satisfy the requirements of an HPC kind even if the boundary separating the affected individuals from the unaffected was fundamentally ambiguous and the affected individuals fell on a gradient of prototypicality (Haslam 2014, p. 18). Partly for this reason, philosophers of psychiatry increasingly endorse Boyd’s as the appropriate concept of kindhood for psychiatric categories (Beebe and Sabbarton-Leary 2010; Kendler et al. 2011).

Along these lines, Samuels (2009) provides the first in-depth discussion of the natural kind status of delusion in particular. He argues for the view that delusion is a natural kind in the liberal HPC sense by skillfully answering various objections to this view and drawing positive morals from them. These objections focus on three characteristics of delusion that may be viewed as flying in the face of its natural kind status: the alleged continuity of delusion with normal experience (van Os et al. 2009); the causal, neural, and cognitive heterogeneity of delusion (Freeman and Garety 2006); and the mind-dependence of delusion as a kind (Murphy 2006). As I am confident that Samuels successfully deals with the first two groups of objections, I will not go over these here, but will confine myself to the mind-dependence objections which, I think, merit further discussion. In the next section, I will set the stage for the discussion of the mind-dependence objections by presenting a model of our intuitive detection and attribution of mental disorder, and an extension of this model that aims at accounting for the detection and attribution of delusion in particular.

2. FOLK PSYCHIATRY AND FOLK EPISTEMOLOGY

2.1 THE DETECTION AND ATTRIBUTION OF MENTAL DISORDER

How do people detect and attribute mental disorder? How do culture-specific models of dysfunction influence these processes? And how do pan-specific features of human minds influence cultural models of detection and attribution? As Pascal Boyer (2011) notes, the actual cognitive processes engaged in when people think about mental disorder have eluded empirical research. He attributes this to the fact that such processes fall between the domains of two well-established disciplines, namely, cross-cultural psychiatry (which focuses on the cultural variation of disorders themselves) and anthropological ethnopsychiatry (which focuses on cultural models of sanity and madness). Recently, however, Haslam and colleagues have, in a series of theoretical and empirical papers, developed a social-cognitive model of laypeople’s thinking about mental disorder—what they dub *folk psychiatry*—which shows promise as an organizing framework for a field that has lacked a clear theoretical basis.

Haslam’s folk psychiatry model specifies four dimensions along which laypeople conceptualize mental disorders: *pathologizing*, that is, the extent to

which the observed behavior is construed as abnormal or deviant, mainly on the basis of rarity, and as a result of the failure to explain the behavior; *moralizing*, the extent to which the observed behavior is under the subject's control and to which individuals are morally accountable for their abnormality; *medicalizing*, the extent to which the observed behavior has a somatic basis and is the direct result of an underlying organic condition; and *psychologizing*, the extent to which the observed behavior has a mental, non-intentional basis, and is the direct result of a psychological dysfunction which shifts the explanatory focus toward causes, not reasons, undermining moral judgment (Haslam, 2003, 2005; Haslam, Ban, and Kaufmann 2007).

Empirical support for the folk psychiatry model comes from a series of studies in which participants rate descriptions of mental disorders and other conditions on a number of items that assess features of the model. In the first study of this sort, Nick Haslam and Cezar Giosan (2002) interviewed American undergraduates who had no formal education in abnormal psychology. They were given the task of reading paragraph-length descriptions of 68 conditions, 47 of which corresponded to DSM-IV mental disorders. They were then asked to judge if the conditions were mental disorders and to rate them on 15 items addressing components of the concept of mental disorder proposed by several theorists. The authors found that American lay understandings of 'mental disorder' showed moderate convergence with the DSM-IV concept of mental disorder. Then, in a follow-up study, Cesar Giosan, Viviane Glovsky, and Nick Haslam (2001) replicated the pilot study in student samples from Brazil and Romania using an identical research design and carefully translated versions of the original questionnaire. The most interesting departure from the American understanding of mental disorder was found among Brazilian participants, who did not represent moralizing and medicalizing as polar opposites, placing them on separate factors and thereby justifying the distinctness and irreducibility of these dimensions.

Besides mapping stable understandings of abnormality within and across cultures, the folk psychiatry model also illuminates shifts in these understandings. Since they found earlier that North American understandings of mental disorders tend to be more psychologized or "internalistic" than those of Brazilians, Glovsky and Haslam (2003) predicted that the longer the period of acculturation of Brazilian citizens living in the United States, the more psychologized their understandings of disorders would be compared to their less acculturated compatriots. Consistent with this prediction, more acculturated participants judged a larger proportion of the conditions to be mental disorders. Importantly, they also understood these conditions more as manifestations of emotional distress and intrapsychic dysfunction and showed a stronger tendency both to understand disorder as a violation of social expectations and to pathologize behavior in excess ('acting out'). Therefore, the concept of 'distúrbio mental' they once shared with their Brazilian peers broadened and took on a more psychologizing cast among more "Americanized" Brazilian participants.

Note, however, that while these studies and the theoretical framework that emerges from them provide an elegant illustration of the cognitive processes of

intuitive detection at work, they do not address the equally important why and how questions about our intuitive detection of mental disorder—namely, *why* and *how* intuitive folk psychiatries emerge. Toward that end, Boyer forges a cognitive model that builds on the evidence provided by Haslam and colleagues, as well as on observations about the causal connections between pathology, cultural context, typical manifestations, popular categorization, and scholarly description. In the first stage of Boyer’s account, dysfunction triggers behaviors, only some of which are detectable as violations of folk psychology—that is, the shared set of assumptions that are the basis of our ability to describe, interpret, and predict each other’s behavior by attributing beliefs, desires, hopes, feelings, and other familiar mental states. (The ones that are not bounce off intuitive detection.) Importantly, sometimes causes other than dysfunction will trigger behaviors that will be interpreted as violations caused by dysfunction, and in these instances detection will have gone wrong. Detection of unexpected behavior will trigger explanatory causal models for the behavior, not all of which make it through cycles of acquisition and communication (unsuccessful models bounce off transmission). Finally, frequently activated models may have feedback effects. These affect the models themselves through the work of transmission biases whereby people are more likely to adopt and transmit representations that are already widespread (Boyd and Richerson 1985). Moreover, they affect people’s behaviors when subjects of classification become aware of being so classified. Such changes, in turn, may lead to revisions in the initial descriptions of mental disorders (Hacking 1995).

For our purposes, what is especially important are the first stages in Boyer’s account, which, in short, boil down to the claim that our intuitive detection of mental disorder involves judging that certain kinds of behavior are so different from our expectations that they are taken as evidence that the mental systems that produce them are dysfunctional. These are mental dispositions that form part of our shared cognitive architecture (Sperber 1996). But just as ‘narratives, scholarship, etiquette, politics, cuisine, musical traditions or religious rituals’ (Boyer 2011, p. 112) are culture-specific, the manifestations of these dispositions to attribute dysfunction will—by deriving from the sets of mental representations that constitute the models of what is wrong with people’s behavior within specific contexts—also be culture-specific. While Boyer’s theory is not a theory of mental illness, but a theory of its attribution, his idea of mental disorder as a defeater of folk psychology may have an important impact on the project of uncovering natural psychiatric kinds, including the project of vindicating the natural kind status of delusion.

2.2 THE FOLK EPISTEMOLOGY OF DELUSION

In the context of a discussion about what he calls the ‘counterintuitive biology’ inherent in some religious and magical concepts, Boyer (2001) considers Wendy James’s account of ‘ebony divination,’ a practice of the Uduk-speaking peoples that she encountered while carrying out fieldwork in the borderlands of Sudan’s frontier with Ethiopia in the 1960s. The Uduk report that ebony trees can

eavesdrop on conversations and that they ‘know of the actions of the *arum* [souls, spirits, including people who were not given a proper burial] and of *dhatu* (witches) and other sources of psychic activity’ (James 1988, p. 303). According to James, diviners perform oracular consultation by burning ebony wood as a form of seeking personal healing and keeping foreign gods at bay. During the consultation, the ebony stick will produce specific smudges in the water which indicate not only the nature of the problem at hand but also a solution.

In contrast, consider the following case described by Murphy:

Ed was sleeping rough, and heard (or, had the experience of) a tree in a park tell him that the park was a good place to stay. So Ed settled down for the night in the park. But a little later, the sprinklers in the park erupted and Ed was drenched. Thereupon Ed heard the tree tell him that he (the tree) was very sorry: trees like to be watered, and the tree had not understood that Ed would not appreciate a good soaking. Ed accepted the tree’s apology and went on his way. (2013, p. 118)

Why is it intuitive to attribute dysfunction in Ed’s case, but not in the Uduk’s case? In addition to characterizing delusion as a false belief based on incorrect inference that is firmly held despite what almost everyone else believes and despite being confronted by evidence to the contrary, the DSM’s definition continues in the following way: ‘The belief is not ordinarily accepted by other members of the person’s culture or subculture (i.e., it is not an article of religious faith)’ (American Psychiatric Association 2013, p. 819). At first glance, this cultural exemption clause may appear to be a highly arbitrary, relativistic, and even unscientific addition. As epistemology does not generally regard widespread cultural endorsement as a form of justification, this sort of exceptionalism has often been dismissed as unwarranted and question-begging (Radden 2011, p. 101).

But the cultural exemption clause in the definition of delusion encodes the fact that *other causes* would be assumed rather than dysfunction in the latter case. Uduk people who believe that trees can hear conversations are members of a culture wherein trees are believed to have counterintuitive biological characteristics, whereas Ed is not. According to Samuels’s interpretation of cultural exemption, in the case of the Uduk the causes of what might seem aberrant behavior for outsiders will, on close inspection, have to do with testimony: when we acknowledge that the belief that trees have counterintuitive biological characteristics is part of the Uduk culture and is acquired through testimony, the need to attribute dysfunction vanishes. In short, testimony *explains* the acquisition of strange beliefs. But what about Ed’s case? Should we conversely interpret the intuitive pull to attribute dysfunction to him as being a result of Ed’s not having the epistemic warrant that the Uduk have through testimony? As much as Samuels’s observations about testimony make sense of cultural exemption in the detection and attribution of mental disorder, the converse interpretation in Ed’s case makes the treatment of delusions implausible, as lack of testimonial warrant is too narrow a rationale to account for our intuitive attribution of delusion. For this reason, Murphy (2014, p. 114–

5) argues that to explain the attribution of delusion we should think more broadly about reasoning, going beyond testimony.

In consonance with Boyer's cognitive account of detection and attribution, Ed's traffic with trees is readily taken as evidence of mental dysfunction in the absence of cultural exemption. Notwithstanding the fact that the description of Ed's experience is one of hallucination, the fact that he accepts this experience as true, inferring that trees can talk and letting his behavior be guided by this conviction, supports the attribution of an accompanying delusion. Murphy (2012, 2013, 2014) applies Boyer's framework to the case of delusion by hypothesizing that the psychiatric concept of delusion grows out of a widespread human tendency, which Boyer accounts for via cognitive science, to attribute mental disorder in cases where someone's behavior fails to accord with folk-psychological assumptions about how the mind works. More specifically, Murphy proposes that our practices of attribution suggest that a delusion is a belief that is acquired through a process that does not fit our folk theories of belief acquisition—which he dubs *folk epistemology*. Unlike the DSM definition, then, Murphy suggests that what is crucial to demarcating delusion from other kinds of aberrant beliefs is not the end product of reasoning but the process by which these beliefs are formed.

What is conceptually basic about delusion is the perversion of normal mechanisms of belief acquisition and revision, not just the weird beliefs that one ends up with through that perverted changing of one's mind. "Normal" here does not mean "according to our best scientific theory." It means that folk psychology, broadly construed, endorses some avenues of belief formation and rejects others. Delusional people are people who are hooked up to the world in ways that ... folk epistemology says are weird, in the sense of falling outside normal human expectations about other people's psychology. The weirdness of the ensuing belief is (defeasible) evidence for the abnormality of their reasoning mechanisms, but the weirdness itself is not the conceptually crucial element. (2014, p. 115)

Thus, what makes delusions distinctive is not that they violate epistemic norms, per se. Instead, our folk-epistemological *expectations* are violated. All manner of beliefs that violate epistemic norms are part of our folk-epistemological expectations and can be accounted for by our folk-epistemological resources which, Murphy (2012, p. 22) elucidates, do not just include folk psychology in the narrow sense of theory of mind, but also beliefs and expectations about the role of "hot" cognition and personal interests in the formation and maintenance of belief, as well as the role of culture in shaping people's assumptions about what counts as legitimate evidence. In the case of self-deception, for example, though the belief is formed and maintained in the face of contradictory evidence, we as interpreters do not run out of explanatory resources and can readily come up with an explanation of how and why the belief came about. In other words, what is distinctive about delusion is the "explanatory gap" created by its observation, and closed by its attribution.

3. ASSESSING THE MIND-DEPENDENCE OF DELUSION

How does Murphy's Boyer-inspired account of delusion attribution impact the status of delusion as a natural kind? Unlike biological taxa which, as we have seen, are prime examples of property clusters held together by homeostatic causal mechanisms, delusion (as well as other psychiatric categories) appear to be mind-dependent (or response-dependent) in ways that put pressure on even the most liberal sense of natural-kindhood.

3.1 DELUSION AS A FOLK-PSYCHOLOGICAL KIND

The first mind-dependence objection one may extract from the discussion of the attribution of delusion simply states that delusion is not a natural kind because it is an artifact of our folk psychology. As Murphy claims, 'whether or not something is a delusion is a matter of how it strikes us, and that depends on how well it comports with our understanding of what people are like, both in general terms and within our culture' (2006, p. 180). Note, however, that even if we follow Samuels and derive such an objection from Murphy's claim that delusion is a matter of how it strikes us, this objection could not be derived from the mere fact that delusions are a part of our folk conception of the world, since there is no immediate incompatibility between the naturalness of a kind and the fact that it maps onto our folk conceptions.

As Samuels notes, water is plausibly a natural kind, though 'water' and the concept it expresses are also part of our folk conceptions. Though one may have affinities for eliminativism concerning some of our folk concepts, there is, on the other hand, no principled reason to deny that at least some of our folk concepts do pick out natural kinds. What the present objection hinges on is the premise, attributed by Samuels to Murphy, that *what it is to be a delusion* is determined by how it strikes us. That is, the premise that all there is to being a delusion is to be a certain kind of response-dependent property. As we have seen, Samuels alludes to Page (2006)'s notion of individutive independence—the sense in which a class of things is circumscribed by boundaries that are totally independent of our taxonomic practices—as the relevant sense in which natural kinds must be response-independent. So the objection at hand can be seen as likening the individuation of abnormal psychological conditions into delusions to the individuation of the night sky into constellations: just as the existence of constellations is parasitic on the way we choose to categorize things, so does the existence of delusions. In other words, the task for those who wish to argue that delusion is a natural kind consists in showing that delusion as a kind is more akin to stars than to constellations.

Samuels's answer to the response-dependence objection consists in arguing that it conflates the metaphysics of delusion with its epistemology: 'The relevant metaphysical issue concerns the nature of delusions: roughly, what is it to be a delusion. The relevant epistemic question concerns the evidential basis for our judgements about delusion: roughly, the sorts of evidence we invoke in judging that someone is deluded' (2009, p. 68–69). Samuels concedes that Murphy gets the epistemology of delusion right, and that not only everyday judgments about

which mental states are delusions are made on the basis of commonsense psychological considerations, but the judgements of clinicians who diagnose delusions are also largely dependent on the same folk conceptions. Samuels's point, then, is that the fact that the detection and attribution of delusion is a matter of how it strikes us does not show that what it is to be a delusion is *exhausted* by how things strike us and, consequently, there is still a possibility that, in this case, our folk conception will be vindicated by, and map onto, a scientific understanding of delusion—what Murphy (2014, p. 119) aptly calls *the vindication project*.

3.2 THE CULTURAL RELATIVITY OF DELUSION

The second mind-dependence objection to which Samuels refers is that which states that delusion is not a natural kind because delusion is context-sensitive. In fact, there are two senses in which delusion may be said to be culturally relative. The first sense expands on what has been just discussed, namely, the fact that the *attribution* of delusion derives from our folk conception of what is and isn't a healthy or normal state of mind. Whereas the previous objection concerns an allegedly universal feature of human folk psychology, a new objection may hinge on the claim that the attribution of delusion will also depend on what is considered a healthy or normal state of mind within one's cultural context, encoded in the cultural exemption clause in the definition of delusion given in the DSM-5. The clause makes sense of the intuition that the delusional individual *stands alone* in some sense (Leeser and O'Donohue 1999, p. 692). The intuitive character of the cultural exceptionalism clause can be seen by contemplating what we would judge as strange and even irrational beliefs which are nevertheless commonplace in cultures other than our own. For example, consider the following entry in Dan Sperber's field diary, from the period he conducted ethnographic fieldwork among the Dorze people of Southern Ethiopia between 1968 and 1974:

Saturday morning old Filate came to see me in a state of great excitement:

“Three times I came to see you, and you weren't there!”

“I was away in Konso.”

“I know. I was angry. I was glad. Do you want to do something?” “What?”

“Keep quiet! If you do it, God will be pleased, the Government will be pleased. So?”

“Well, if it is a good thing and if I can do it, I shall do it.”

“I have talked to no one about it: will you kill it?”

“Kill? Kill what?”

“Its heart is made of gold, it has one horn on the nape of its neck.

It is golden all over. It does not live far, two days' walk at most. If you kill it, you will become a great man!”

And so on . . . It turns out Filate wants me to kill a dragon. He is to come back this afternoon with someone who has seen it, and they will tell me more . . . (1982, p. 35)

Commenting on this entry, Sperber goes on to express respect and affection for his Ethiopian friend. He is confident that the man was not senile at the time of the unusual request and, moreover, that he was too poor to drink. Consequently, Sperber is faced with a variation of a question that, undoubtedly, all of us ask ourselves of someone else at some point: how could a sound person believe *that*? ‘That’ being, in this case, that dragons exist, not “once upon a time,” but there and then, within walking distance. What if Sperber had expressed doubts that such an animal even exists? What if he had pressed his friend on the issue of the dragon’s heart being made of gold and the apparent impossibility of a gold heart *beating*? Sperber concludes that his friend was ‘merely quoting what people who had killed these animals were reported to have said, and they knew better than any of us’ (1982, p. 61). In line with Sperber’s explanation, Samuels (2009, pp. 69–70) argues that the cultural relativity of delusions tracks precisely the insensitivity of delusions to testimony—an important source of epistemic warrant and epistemic defeat. Because it is normal for one to form and maintain beliefs based on the testimony of peers and authorities from one’s culture or subculture, resistance to testimony is viewed as a sign that something is wrong. And because one’s source of testimony varies with one’s culture and subculture, the cultural exemption clause is a necessary measure to avoid the hasty judgment that culture-bound beliefs are necessarily irrational and possibly even the product of pre-rational mental processes (Sperber 1980). However, so long as the resistance to testimony that characterizes delusion is culturally *invariant*, the fact that delusions are resistant to testimony does not suffice to show that delusion is a response-dependent property and, thus, cannot be used to successfully object to the natural kind status of delusion.

The second sense in which delusion may be said to be culturally relative derives from the fact that the *content* of delusions is highly sensitive to social and cultural context. So, for example, Masato Tateyama and colleagues (1998) compared the schizophrenic delusions of 324 inpatients in Japan, 101 in Austria, and 150 in Germany, and found that themes of persecutory delusion (i.e., delusions of poisoning) and religious themes of guilt/sin were conspicuous in Europe, while amorphous delusions of reference (i.e. ‘being slandered’) were predominant in Japan. Another study conducted by Thomas Stompe and colleagues (1999) compared the schizophrenic delusions of 126 Austrian and 108 Pakistani patients, finding significantly higher frequencies of grandiose and religious delusions in Austrian patients, and persecutory delusions with political themes among male Pakistani patients. To these observations may be added the existence of culture-bound syndromes whose expression includes culture-specific symptoms, as in *koro*, most prevalent among Chinese ethnic groups, in which an individual claims that his or her genitals are retracting and will disappear (Chowdhury 1996).

Time is also a factor. Changes within one and the same culture have an impact on the diachronic variability of delusional content, as Borut Škodlar and colleagues (2008) have found in a study of admission records of patients with schizophrenia in Slovenia from 1881 to 2000. The recent emergence of the so-called Truman Show delusion attests to the same fact—patients with ‘Truman

signs' claim that their lives are staged plays or reality television shows, as with the protagonist of the 1998 film *The Truman Show* (Fusar-Poli et al. 2008; Gold and Gold 2012). However, though the kinds of variability discussed above may suggest that delusion is response-dependent to the extent that what is a delusion depends on what beliefs are socially prevalent at a certain point in time, Samuels (2009, p. 69) notes that what the sensitivity of delusions to social context shows is only that the nature of delusion, as Karl Jaspers (1913) long before observed, cannot be characterized, but can at best only be classified, in terms of its contents.

3.3 THE VINDICATION PROJECT

If Boyer and Murphy are correct, then the science of delusion is inextricably tied with its intuitive detection. Psychiatric elaborations of folk psychology give rise to the clinical concept of delusion, the extension of which is then subdivided according to surface features, most prominent among these its content (i.e., what it is about). But can delusion, being rooted in folk psychology, play the role of regimenting scientific inquiry?

By defending that delusion is a natural kind in the HPC sense, Samuels answers positively and wagers that scientific psychiatry will vindicate the folk concept of delusion—that is, if Samuels is correct, the folk concept of delusion picks out a causal signature that, once uncovered, will vindicate the reliability of this concept and show that delusion is, in fact, a homeostatic property cluster. Once the causal mechanisms that make the properties of delusion co-occur are discovered, causal classification may result in many current subtypes of delusion being excluded from its extension. But because the HPC conception of natural kindhood does not mandate that natural kinds have category essences or category boundaries, it is likely that a mature science of delusion informed by its causal mechanisms will not be able to give a simple yes or no answer to every question of the form 'Is X a delusion?'. Of more practical importance, however, is the fact that a causal understanding of the underlying mechanisms would suffice to yield powerful inductive generalizations regarding diagnosis, prevention, and management.

But how does the vindication project fare in view of the mind-dependence of the folk concept of delusion? As Samuels notes, this only hurts the chances of delusion being an HPC kind if we conflate the metaphysics and epistemology of delusion. As we have seen, Samuels argues that attention to the fact that our concept of delusion is a part of our folk psychology that has been incorporated into scientific psychology and psychiatry is not enough to show that it is not a natural kind: the folk-psychological kind may well track an underlying natural kind. Samuels (2009, p. 69) notes that, to support the mind-dependence objection, it would be necessary to show that in the case of delusion the metaphysical issues about the nature of the kind and the epistemic issues about how we know about instances of the kind *should* be collapsed. Showing that the clinical concept is built on folk conceptions of normality is not enough. Importantly, however, Samuels does not establish that delusion *is* a natural kind. In fact, he could not have established this on the basis of *a priori* speculation

alone, as establishing natural kindhood is ultimately a matter of investigating the causal basis of the homeostasis of property clusters (assuming the HPC model). Samuels does skillfully argue against various objections to the status of delusion as a homeostatic property cluster, some of which I have discussed above. In doing so, Samuels establishes something very important, namely, that these objections are not sufficient to exclude the possibility that delusion is a natural kind. So what we are left with after Samuels's arguments is that the natural kind status of delusion is still an open question, i.e. that delusion is *possibly* a natural kind.

Although the argument from mind-dependence that derives from accepting the application of Boyer's theory to delusion is not enough to rule out the possibility that delusion is a natural kind, it does make Samuels's thesis implausible and gives him the burden of proof. This implausibility can be better seen if we compare generic folk kinds and generic scientific kinds. If Samuels is right, delusion would be a *generic* natural kind. Just like the kind metal subsumes many different subordinate kinds such as gold, copper, and magnesium, delusion will subsume subtypes which would themselves also be natural kinds. But Samuels's optimism regarding the vindication project is hardly justified by the observation of other generic folk concepts and how they relate to their scientific counterparts. For instance, what the folk concept of metal seemingly picks out is not a causal signature, but, as Murphy (2014, p. 121) notes, a variety of properties that directly relate to our interests, properties like being shiny, being malleable, etc., rather than a chemical element whose atoms readily lose electrons to form positive ions, etc. Likewise, the folk concept of lily, as Dupré (1981, p. 74) points out, does not accurately map onto the biological concept of lily, which includes garlands and onions, but is used to refer exclusively to a type of flower. If delusion picks out properties that relate to our interests, like being *weird* to varying degrees, then the burden of proof falls squarely on Samuels with respect to the likelihood of vindication.

Furthermore, as investigation into the causes of delusion is still in early stages, accepting the view that delusion constitutes an HPC kind is as much a "black-box" approach as Wakefield's, only more modest in its ambition. I have argued that as an ontological commitment, this approach is weak. As a *methodological* commitment, on the other hand—and this is the sense in which Kendler and colleagues (2011) seem to accept that psychiatric categories in general are HPC kinds—there is still a case for viewing delusion as a generic natural kind with an eye toward progress in scientific psychiatry. Bearing in mind that what we are authorized to commit to (ontologically) at this moment is that delusion is a *practical* kind—as this coheres both with our knowledge of delusion in the clinic as well as with our best theory of detection and attribution of mental disorder (and delusion in particular)—if the possibility of natural kindhood is still open, assuming natural kindhood is a sound methodology inasmuch as it offers a way toward progress in causal classification. However, I maintain that this is neither the only, nor the best way toward progress.

Even if the *folk* concept of metal is not appropriate to play the role of regimenting scientific inquiry, chemistry did eventually arrive at the natural kind metal and many subspecies of our folk concept of metal, such as gold, silver,

copper, etc. also turned out to be natural kinds. In this manner, despite delusion being a folk concept not so far mapped onto a rigorous scientific concept, many subtypes of delusion already recognized, such as clear-cut cases of monothematic delusions following brain damage (e.g. Capgras, mirrored-self misidentification, somatoparaphrenia, etc.), might still turn out to be natural kinds which are thrown in with similar conditions that strike us as weird into the set of phenomena described folk-psychologically (and clinically) as delusions. Our focus should be on uncovering the causal mechanisms underlying specific kinds of delusion rather than trying to impose a general causal explanation on a ragbag of different abnormalities that may or may not actually be of the same kind. Thus, I suggest a compromise between Zachar's (2000) earlier work and Samuels's (2009) defense of delusion as an HPC kind, drawing on Murphy's (2014) observations: delusion, as a kind rooted in folk psychology, is probably a practical kind, and it probably does not pick out a universal causal signature that makes the whole category be a natural kind, but it probably does pick out many subspecies which are themselves natural kinds. Hence:

Hypothesis: Delusion is not a natural kind, but some delusions are.

So if the question were 'Is Capgras delusion a natural kind?' or 'Is somatoparaphrenia a natural kind?' being that these are stable clusters of properties with recognizably homogeneous neurological causes and which are not the product of generic folk intuitions but of rigorous clinical observation and investigation, the case for their natural kindhood would be much stronger and plausible. Thus, I suggest that the way to progress in the science of delusion lies in trying to vindicate the natural kind status of *subspecies* of delusions through the study of the causal mechanisms that make the relevant properties occur homeostatically, and not in trying to find a shared causal basis for every phenomena that we call delusion assuming beforehand that such a shared causal basis is present. After the investigation into the causal mechanisms is done with multiple subtypes of delusion, a causal account of delusion in general will no doubt progressively suggest itself. But the set of delusion subtypes that will be found to share causal mechanisms in the sense that would authorize us to abstract from them a generic natural kind will be a *subset* of the set of all delusions—a set the intension of which depends on context-dependent folk-psychological intuitions and, hence, membership in such a set is tied to surface features (symptoms, not causes) detected with the tools of folk psychology.

CONCLUSION

In the preceding sections, I have attempted to elucidate some of the difficulties inherent in trying to claim that delusion is a natural kind. After delineating five different senses of kindhood and introducing a non-essentialist approach to natural kindhood—the HPC model—I have drawn on a cognitive model of the intuitive detection and attribution of mental disorder and its application to the case of delusion to flesh out the fact that the clinical category of delusion is rooted in folk-psychological expectations. Finally, being that the folk-

psychological status of delusion does not immediately remove the possibility of this kind being vindicated as natural by scientific investigation, I have questioned the vindication project and formulated a working hypothesis that I claim is both ontologically and methodologically more sound. My hypothesis is that along with the general category of delusion, some delusions will be confined to practical kindhood, perhaps along with the bulk of mental symptoms and disorders, while some will turn out to be objective distinctions in nature. Importantly, this hypothesis and methodological suggestion bypasses what Samuels calls the *unity problem*: if many different subtypes of mechanism are responsible for delusions, why treat delusions *as such* as a natural kind? According to him, it must be because these mechanisms are themselves of the same kind. What I have tried to show in this chapter is that this is an improbable scenario. Assuming that a variety of mechanisms make subtypes of delusion subtypes of some general mechanism as opposed to a heterogeneous collection of different mechanisms the products of which share surface features is not only unwarranted, but methodologically flawed.

THE ROLE OF *OSTENSIVE TRAINING* IN NATURAL LANGUAGE ACQUISITION

JULIANO DO CARMO

1. DIFFERENT PERSPECTIVES ON THE PROCESS OF LANGUAGE ACQUISITION

The process of natural language acquisition is a widely discussed theme among linguists, philosophers, psychologists and neuroscientists, mainly due to the controversy regarding the explanations offered for the processes of acquisition of the abilities needed for learning and developing language. The fact that learners in a “verbal environment,” even if they lack formal training, generally learn to speak and understand the expressions of language with no great difficulties has been leading many theoreticians to think that there must necessarily be some previous syntactic or semantic structure that allows them to understand and use natural language. As we will see, there is no need to presuppose “innate categories” or even a “private language of thought” to explain the process of natural language acquisition¹.

Noam Chomsky’s innatist (or nativist) argument exposed in *Aspects of the Theory of Syntax* in 1965 aims to show that language learning cannot be dependent on a simple linguistic input, since this input would not be capable of developing the tools needed for the potentially infinite processing of sentences of a natural language in the learner. Chomsky’s strategy is to posit that there are some innate and universal biological grammatical categories (once they would be supposedly present in all the known natural languages), such as nouns and verbs, that enable the complete development of language in the learners. Moreover, the learners have a kind of natural (instinctive) disposition to associate nouns and verbs and, thus, to form complete and significant sentences. As is well-known, the success of Chomsky’s strategy depends on the presupposition of a Universal Grammar (innate biological categories). However, the idea of a Universal Grammar has been strongly refused by the theoreticians of cognitive psychology, mainly with regard to the supposed basic categories (nouns and verbs), since, from a biological and evolutionary point of view, the motives for these same categories being selected are not clear. Innatism should be able to explain how the “language faculty,” containing innate representations of the Universal Grammar, has arisen in the human mind.² Until now, however, there is no consensus about how the knowledge of the Universal

¹ I would like to thank those with whom I had the opportunity to discuss (in many events) some of the ideas developed in this article, in particular my dear colleagues Carlos Miraglia and Flávia Carvalho Chagas for the excellent remarks, critiques and suggestions.

² See: COWIE, F. *Innateness and Language*. Stanford Encyclopedia of Philosophy, 2008.

Grammar (or of its categories) may have evolved out of any adaptive element in our ancestors, since linguistic ability does not seem to offer any reproductive advantage.³ Instead of presupposing specific innate mechanisms for language processing (from an intellectualist point of view), maybe it could be the case of presupposing that the learners of a natural language use cognitive principles of general learning. In addition, Chomsky's innatism presupposes a mechanism of language processing that is excessively precocious for learners, since it assumes that learners have abilities that are fully developed in the competent users of the language.

On the other hand, Jerry Fodor's apriorist thesis, exposed in *The Language of Thought* (1975), is based on the idea that it is not possible to explain the process of learning a natural language without presupposing a "private language of thought." Fodor advocates that, for learning a natural language, human beings need to have access to the representations of certain conceptual propositions, that is, they should be able to manipulate certain concepts.⁴ Natural languages are public and controlled by equally public criteria. However, in the moment of acquiring a natural language, human beings have no access to a natural language. Therefore, Fodor posits that the concepts needed for learning to occur cannot be controlled by public criteria and, thus, the previous manipulation of certain concepts should be controlled by regularities that are not public. Fodor's strategy (concerning the explanatory need of the supposed private language of thought) seems to work only insofar as the computational model of language is true, and this seems to be an excessively high price to pay for an explanation of the process of language acquisition. In addition, Fodor's intellectualism seems to presuppose that speaking a language (as an intelligent action) is an action necessarily preceded or directed by a previous theoretical operation, such as the intellectual action of considering a regulative proposition. If each and every intelligent action presupposes a previous theoretical operation, then it seems really difficult to prevent the infinite regress of theoretical operations.⁵

Skinner's behaviorism, whose bases are found in *The Verbal Behavior* (1957), has been the most important rival perspective to innatism and apriorism, once it seeks to explain the process of language acquisition through empirical, actual and publically observable events. The basic idea is that the acquisition and the development of a natural language occur through the learner's exposition to influences of the verbal behavior of the competent users of the language. Insofar as there are no innate abilities, it is the verbal environment that enables the association of words and meanings.⁶ A verbal environment is the product of the constant

³ See: LEWONTIN, R. *The Evolution of Cognition: Questions we will Never Answer*, In R. Sternberg and D. Scarborough (eds.). *An Invitation to Cognitive Science*. Cambridge: MIT Press, 1998.

⁴ See: FODOR, J. *The Language of Thought*. New York: Thomas Y. Crowell Company, 1975.

⁵ RYLE, G. *The Concept of Mind*. New York: Routledge, 1949. p. 31.

⁶ Skinner's behaviorism aims to explain language acquisition in terms of the subject's previous history, current circumstances and nothing more. According to Stephen Winokur, the idea is to show, from an analysis of the speech behavior, that language acquisition is linked to many variables whereof verbal behavior is a function. Skinner's psychology,

interchange among competent users of the language.⁷ The behaviorist strategy consists in advocating that the correct utterances of a learner are positively reinforced when he or she understands the communicative value of words and sentences. Language is fully acquired through an operative conditioning where the learners perceive certain regularities within the verbal environment through positively reinforced behaviors.

The variables whereof behavior is a function are sought in empirical, actual and publically observable behaviors. The organism's present and past states are observed (deprivation), its genetic constitution (a human being or a chimpanzee may acquire verbal behavior, but a pigeon may not), and the present and past states of the environment (stimuli); and all of them are actual, physical and intersubjectively observable things. The covariations of changes in these empirical variables and the changes in behavior may be described. The description of functional relationships among empirical variables constitutes the type of causal consideration that interests us.

What renders the classic versions of behaviorism not very attractive, however, to counter innatism and apriorism is the supposed difficulty in explaining the way the subjects become autonomous in relation to conditioning. Skinner also recognizes that the problem of determining how verbal environment may have arisen from a non-verbal source is the target of many speculations and that the answers offered to this problem are no more than inferences to the best explanation.⁹ Moreover, the explanations presented to make this passage most of the times seem excessively artificial. Yet, these difficulties do not fully disable an anti-intellectualist position on the process of natural language acquisition, but demand that the theoretician provide a reasonable explanation for the process as a whole, as will be seen.

Studies on encephalic processes in language acquisition have been leading to considerable advancements in relation to the description of areas involved or

based on the notion of stimulus-response, understands behavior as the central point of investigation. The movements produced by the muscles of animals should be explained instead of ideas, meanings, desires, intentions, expectations or any type of hypothetical psychological mechanism. The idea, in other words, is to highlight that verbal behavior is also a kind of behavior (not too different from non-verbal behavior). According to Winokur, to make a noise with the mouth (to speak, for instance) is a behavior as much as riding a bicycle. Cf.: WINOKUR, S. *A primer of verbal behavior: an operant view*. Englewood Cliffs: Prentice-Hall, 1976. 1-3. Fodor harshly criticizes this model of thinking, mainly concerning the notion of "proximal stimulus" as a proximal representation of the "distal stimulus." Insofar as there is no internal representation without an internal language, Fodor believes that behaviorism is not a reasonable way of explaining the process of natural language acquisition. Cf. FODOR, J. *The Language of Thought*. New York: Thomas Y. Crowell Company, 1975, pp. 55-65.

⁷ SKINNER, B. F. *Verbal Behavior*. New York: Appleton-Century-Crofts, 1957. p. 461.

⁸ WINOKUR, S. *A primer of verbal behavior: an operant view*. Englewood Cliffs: Prentice-Hall, 1976. 1-3.

⁹ SKINNER, B. F. *Verbal Behavior*. New York: Appleton-Century-Crofts, 1957. p. 461-5.

activated when we acquire, understand and use language.¹⁰ Neuroscientists in general agree with the thesis that the brain has an innate ability to acquire language. Neural plasticity,¹¹ as an anti-intellectualist element, has been required as an indispensable element for acquiring a language and also for molding most of our acquired behaviors, but empirical studies that show that neural plasticity is a necessary and sufficient element to explain semantic normativity are still lacking, for instance. That is, despite the fact that science aims to offer a detailed description of brain processes involved in the acquisition of the rules of a language, it supposedly cannot adequately explain what obliges or constrains us to use or apply these rules in the same way as we acquire them in the most varied situations or language games¹² (including in situations to which we have not been previously exposed¹³).

The argument that will be developed here presupposes the idea that a natural language demands different abilities from its users and that these abilities¹⁴ are not innate or a priori, but actually totally acquired. In my opinion, there is no doubt that the hypothesis of the neurosciences is correct (that there must exist an innate cerebral “capacity” for language acquisition), but I will argue that although this is a necessary condition, it is definitely not a sufficient condition, since many “abilities” need to be acquired for learning to occur.

¹⁰ Neurosciences aim to explain the supposed universality of language due to the development of specific areas of the brain for language processing (mainly Broca’s and Wernicke’s areas). However, even if most studies and findings of the neurosciences result from studies about partial or complete loss of linguistic abilities (aphasia) due to encephalic lesions, research about the process of first natural language acquisition is scarce. Recently, Patrícia Kuhl has argued in favor of the universality of language in humans, showing, through magnetoencephalography, that children learn a natural language by exposition to speech differently than other animals, but has also highlighted the importance of social interaction to language learning. The results of neuroimaging and encephalic stimulation in general seem consistent with the areas mapped for language processing (inferred from aphasias). However, everything seems to indicate that language acquisition and processing involves much more than a simple relationship between Broca’s and Wernicke’s areas. See: KUHL, P.K. *Learning and Representation in Speech and Language. Current Opinion in Neurobiology*, 4: 812-822, 1994.

¹¹ See: BEAR, M.; CONNORS, B.; PARADISO, M. *Neurociências: Desvendando o Sistema Nervoso*. 3ª Edição. Porto Alegre: Artmed, 2010.

¹² The general idea is that the investigation of the process of language learning is partly empirical and partly philosophical. See: HORWICH, P. *Naturalism, Deflationism and the Relative Priority of Language and Metaphysics*. In: Price, H. *Expressivism, Pragmatism and Representationalism*. New York: Cambridge University Press, 2013; and ENGELLAND, C. *Ostension: Word Learning and the Embodied Mind*. Massachusetts: MIT, 2015.

¹³ I refer here to Kripke’s skeptical argument exposed in “*Wittgenstein on Rules and Private Language*”.

¹⁴ It is important to observe that “capacities” are not identified here with “abilities.” The main difference between them is the fact that, in general, abilities are seen as something learned throughout life whereas capacities are considered something we are born with and may develop. Thus, understanding, for instance, will be seen as an ability (such as running and jumping) while neural plasticity is a capacity (such as motor flexibility and agility).

2. NATURALISM AND NORMATIVITY IN THE PROCESS OF NATURAL LANGUAGE ACQUISITION

The usual way of explaining the process of natural language acquisition is taking the “ostensive teaching” of phrases in the language as starting point. “Ostension” is traditionally considered to be a bodily movement that supposedly manifests our intentionality in relation to the things or events in the world. If gestures and facial expressions are able to reveal intentional states, then ostension really seems to be extremely important for language acquisition, since it can provide the learner with a “pre-linguistic” way of understanding the words and sentences of a natural language.¹⁵ Clearly although ostension is one of the important elements of language acquisition, it is certainly not the only one. The learners should have a sensorial perception of the world and also of the types of things that should be found through it. Engelland¹⁶ posits that a natural language does not demand from its users only semantics, but also syntax, and syntax seems to demand much more than ostensive gestures. In addition, the learners need various abilities and capacities, such as the articulation of words in a meaningful discourse, memory and motivations. Then, ostension is certainly a necessary, but not sufficient condition for language acquisition.

The role of ostension in the process of language acquisition is not always clear in specialized literature. Depending on the theoretical approach, ostension may be seen as the primary or secondary element in the learning process. Innatists and apriorists usually consider ostension as secondary in relation to innate abilities and categories or to the supposed “language of thought” for the development of language. Classical behaviorism and denotational theories of meaning generally consider ostension as primary for language acquisition or learning. In the anti-intellectualist perspective supported here, ostension is primary in developing the necessary abilities to language acquisition and secondary for the development of language learning.

There is an important difference between what has been called “ostensive training”¹⁷ and what is generally conceived of as “ostensive definition.” This distinction was evidenced by Wittgenstein in his *Philosophical Investigations* (1953), mainly in passages aimed at refusing Augustine’s view of language. Recently, however, Meredith Williams recovered this distinction, mainly in *Blind Obedience* (2014), widening the discussion on the process of primitive language acquisition and its interface with the representational theories of meaning and with the neurosciences as well.

¹⁵ Ostension has been a widely discussed theme (at least since Plato) in the history of philosophy and has figured contemporarily as an essential ingredient in the debates on the philosophy of language and mind. Generally speaking, the question is whether intentions may be manifested (or expressed) or can also be inferred; if it is possible to adequately or unequivocally understand an “ostensive gesture.”

¹⁶ ENGELLAND, C. *Ostension: Word Learning and the Embodied Mind*. Massachusetts: MIT, 2015. p. 11.

¹⁷ WILLIAMS, M. *Blind Obedience*. New York: Routledge, 2014. p. 80.

Throughout the 20th century, “ostensive definition” was a widely discussed theme within philosophy, mainly in works by Wittgenstein, Quine and Davidson. It should be noticed that the first paragraph of Wittgenstein’s *Philosophical Investigations* (1953) specifically questions the role of “ostensive definition” in the process of language acquisition (against the denotational view of meaning). Quine harshly criticized the notion of ostensive definition, mainly with regard to the problem of the ambiguity of the “ostensive gesture,” which is evidently related to the problem of “indeterminacy of translation” and of “inscrutability of reference.” Davidson sought to solve the problem of “ostensive definition” through the “triangulation” process (the triangular process of interaction between two subjects and the interaction between each subject and a common set of objects in the world). For our own purposes, the central problem of “ostensive definition” is exactly that of considering it the primary element in the process of language acquisition.

Insofar as the classic behaviorist advocates a non-cognitivist or non-intellectualist perspective, he or she supposes that nothing beyond the ostensive gesture is needed for acquiring the first words of a natural language. I believe that the behaviorist is right in advocating ostension as primary, since it is assumed that nothing beyond training is needed to introduce a learner within the complex universe of language and no intelligent operation related to the process of language acquisition¹⁸ would be explained in terms of previous operations. On the other hand, I believe that the behaviorist is wrong, however, insofar as he or she fails to distinguish between “ostensive training” and “ostensive definition.” In this sense, behaviorism leads to a mistaken image of the “situation of learning,” since it presupposes that the learner has the same abilities that are already fully developed in the competent user of the language.

Analogously, denotational theories of meaning usually propose the thesis that the object denoted by a word fixates the meaning of this word in providing a pattern for the right application of the term to other objects of the same class and in other contexts. In this sense, denotational theory is closely associated to that which became known as “Augustinian view of language,” which posits that adults name individual objects when they point them in the immediate environment of the learner. The learner, as a result of ostension, associates the name uttered by the adult to the object denoted. After making this association, the learner uses this name to

¹⁸ According to Ryle, if the occurrence of any intelligent operation necessarily demands the occurrence of a previous theoretical operation and that that be executed intelligently, then there would be a logical impossibility to anyone who has never entered this circle and, thus, the acquisition of a first natural language would be impossible. If using a natural language is an intelligent action, then it should be presumed that in some moment the acquisition of this natural language was not the result of a previous theoretical operation. Ryle’s objective is showing that some applications of certain rules, principles or norms do not necessarily require a previous consultation of these rules, principles or norms. For Ryle, both idealism and reductionism are responses to an illegitimate question. Both the ‘reduction’ of the material world to mental states and processes and the reduction of mental states to physical processes presuppose the legitimacy of exclusive disjunction: either minds or bodies exist (but not both). It is perfectly correct to say with a certain logical adjustment that there are minds and to say, with another logical adjustment, that there are bodies. Such expressions do not indicate two different types of existence.

any object that belongs to the same class of the originally named object. The fact is that it is not obvious that the simple act of denoting is enough for causing the desired association within the learner's mind and that he or she will make the further generalization in relation to the same objects as the competent user of the language.

The intellectualist posits that the relationship established between word and object is not a simple association, but actually an "ostensive definition."¹⁹ The very effectiveness of ostensive definition depends on the learner's cognitive ability in understanding that the object baptized by the competent user of the language is the exemplar that works as the pattern for the right application of the term uttered. The mistake of denotational theory lies in not perceiving that the situation of learning is not equal among the competent users of the language and the learners. Actually, as we see, the situation of learning is fractioned into two distinct, though complementary, domains: the cognitively impoverished domain of the learner and the cognitively enriched domain of the competent user of the language.²⁰

The mistake of treating ostensive definition as the vehicle to fixate meanings lies in confusing ostensive definition with ostensive training. Ostensive training is used with the beginner learner. Once ostensive training may be characterized as close to stimulus-response conditioning, examples of this training may be described in relative isolation of any background, social convention or costume. This training results in a way of "categorizing" that is not a question of naming and does not presuppose a background and conceptual competence by the learner (...).²¹

According to Meredith Williams, the "learning situation" is also the fundamental element to explain the problem of normativity, that is, the problem of reconciling causal determination of ostensive training with logical or normative determination for the application of acquired rules or norms.²² This is something that behaviorism cannot explain alone. The general idea is that before acquiring a natural language, the learner needs to acquire the necessary abilities to understand a language.²³ The abilities are transmitted in a non-cognitive manner, that is, through examples, expressions of approval, expressions of disapproval, expressions of encouragement and others.²⁴ A competent user of the

¹⁹ WILLIAMS, M. *Blind Obedience*. New York: Routledge, 2014. p. 81.

²⁰ *Idem*, p. 19.

²¹ *Idem*, p. 81.

²² *Idem*, pp. 80-83.

²³ WITTGENSTEIN, L. *Philosophical Investigations*. New York: Blackwell, 2001. §86.

²⁴ According to Williams, Wittgenstein identifies three stages for the acquisition of a natural language: the first is the "stage of configuration," wherein the learner is submitted to a series of examples about what it means "to do the same thing" (following a rule or an arithmetic progression, for instance); the second is the "stage of mastering techniques and practices" and the third is the "stage of certainty." Any use of the language demands the exercise of techniques to follow in a similar way and a shared minimal sense of what is obvious and what is free from doubt. Understanding a language, taking understanding as an acquired ability, is to be able to recognize the verbal behavior of a community and be disposed to follow its rules (even if blindly). See: WILLIAMS, M. *Blind Obedience*. New York: Oxford University Press, 2014.

language is that who possesses the semantic and cognitive abilities to manipulate concepts. As Wittgenstein says in the *Philosophical Investigations*, those who do not have these abilities should be taught through “examples” and “exercises,” through “ostensive training.”

In the course of this teaching I shall shew him the same colours, the same lengths, the same shapes, I shall make him find them and produce them and so on. I shall, for instance, get him to continue an ornamental pattern uniformly when told to do so. — And so to continue progressions. I do it, he does it after me; and I influence by expressions of agreement, rejection, expectation, encouragement I let him go his way, or hold him back; and so on.²⁵

Teaching through examples and repetition through exercises, influence through expressions of approval or disapproval, expressions of expectation and encouragement, letting it continue or preventing it are features of the “ostensive training” for the acquisition of the necessary abilities to learn a language. “Ostensive training” does not necessarily imply previous theoretical operations or the knowledge of true propositions about how to speak a language. Ostensive training, from a non-intellectualist standpoint, aims at transmitting a not-yet developed ability: the ability to understand and manipulate concepts.

The learner may not think properly due to not possessing the basic conceptual requisites, but also because it is only by observing particular examples presented by the instructor that the learner can acquire the sense of what is “to follow in one same way” (“sameness judgments,” in Meredith Williams’ words). The learner acquires the “tools” to act and judge through “examples, practices, and techniques” to participate in language games. Wittgenstein advocated that training in the use of rules demands a background (a scenario) in the same way as training by the practice of naming (only those who can ask the name of things are able to understand an ostensive definition²⁶). There must be a background within which training is done and the ostensive gesture “and so on” is transmitted.

Ostensive training explores the causality based on our perceptual sensibilities in relation to certain objects and properties in the world, whereas “training in normative practices”²⁷ explores the malleability of our behavior and our capacity of offering responses to its validation. The malleability in our behavior is supported by neural plasticity, that is, through the capacity of our brains to develop new synaptic connections among neurons from the person’s experience and behavior and from certain stimuli. Obviously none of these causal propensities concern the normative structure of language, which is governed by *correction patterns*, not *causal laws*. Natural dispositions to acquire certain basic regularities of use²⁸ are explored in initiating the learner in the normative

²⁵ WITTGENSTEIN, L. *Philosophical Investigations*. New York: Blackwell, 2001, §208.

²⁶ *Idem*, §30.

²⁷ WILLIAMS, M. *Normative Naturalism*. International Journal of Philosophical Studies. Nº 18 (3): pp. 355-375, 2011.

²⁸ See: HORWICH, P. *Meaning*. Oxford: Oxford University Press, 1996.

regularities of costumes or practices. Differently from intellectualist positions, in primitive language acquisition, there are no concepts for the learner, not even a grammatical background or articulated environment. The explanations are transmitted through examples, encouragement and discouragement to which the learner seeks to conform *blindly*. Once no expression is applied in itself, there is no logical circle for the explanations provided to the learner.

It is only through the mastery of the competent user of the language that the words and actions of the learner become correct or incorrect (they count as genuine reports or not). The learning situation supposes, on the one hand, the cognitive and normative labor division between the learner's actions and judgments and, on the other hand, the background against which the identity of actions is fixed and the corrections may be performed.²⁹ It is equally important to notice that the learning situation is an initial situation in which the learner must perform "judgments" and be engaged in "reasoning" before being able to do it.

Learners are "calibrated" to judge in conformity with the community, so much so that afterwards they become able to think and act independently from the community. The learner becomes autonomous insofar as he or she accepts the standards sanctioned by the community.³⁰ In this specific sense, we understand the importance of ostensive training and the transmission of "judgments of normative similarity": the acceptance of community standards is blind and unavoidable, since there are no alternatives available. One way of elucidating the stage of calibration is through Wittgenstein's analogy of the standard-meter. There are three fundamental features of the "standard-meter" highlighted by Wittgenstein:

- (1) It is the use of a device (or its "purpose") that makes it widely recognized as a standard-meter;
- (2) There are certain "material properties" that are necessary for an object to work as a one-meter-bar (its rigidity and its capacity to be etched with a scale, for instance);
- (3) "Measurement scales" are needed (centimeter, millimeter, and the meter mark etched in the bar's body).³¹

None of the properties of the one-meter-bar may be considered as inherently representational or normative, since anything used as a meter should have the material properties needed to perform this role. The one-meter bar defines a measurement pattern (a conventional measurement system) due to being used for this purpose. The fact that the metal bar may be used for this purpose is a function of its physical rigidity and of our attributing this role to the bar. It may happen, however, that the material used to build the one-meter-bar is never used for this purpose, that is, the metal in itself measures nothing.

²⁹ WILLIAMS, M. *Blind Obedience*. New York: Oxford University Press, 2014. p. 80.

³⁰ WILLIAMS, M. *Normative Naturalism*. International Journal of Philosophical Studies. N° 18 (3): pp. 355-375, 2011.

³¹ WITTGENSTEIN, L. *Philosophical Investigations*. New York: Blackwell, 2001, §§39-50.

Analogously, we can say that we have three essential features of conventional measurement standards: purpose, material properties and scale. Our sensorial systems and our behavioral malleability (supported by neural plasticity) are the material properties that allow us to measure the world. We are “calibrated” to measure colors, objects, numeric sequences etc. Normative similarity judgments work as the scales that are etched in us for the purpose of measuring many aspects of the world. Due to the acquisition of so many judgments, we measure the world instinctively for that which is obvious and for that which is free from doubt.³²

The most significant difference between the one-meter-bar and our own concerns is the purpose of these standards. The use of the one-meter-bar is a deliberate action while recognizing colors and objects does not involve a deliberate use of ourselves as instruments.³³ We simply see colors. We do not use our bodies for this purpose. The success in making measurements demands that all reach the same measures: where one subject S judges x as something red, it is necessary that other users agree with this judgment. This aspect of sociability of the pattern is the empirical rigidity demanded to perform measurements. The theoreticians of the community concerning what does it mean to follow a rule are wrong exactly in this respect, since “This is red” does not necessarily mean that “the members of the community call it red.” As is highlighted by Williams, taking the fact that “the members of the community call it red” as criterion would be the same as taking the table as the measure of the ruler, and we would no longer be measuring in terms of “meters,” but actually in terms of “tables.”³⁴

The rigidity of intersubjective conformity about what is a standard-meter is the very rigidity of causation. The etchings on the metal-bar are the results of the causes operating on the metal’s surface. The fixedness of the etchings is as important as the rigidity of the bar itself. Analogously, the “judgments about what is to follow the same way” (the sameness judgments) that are constitutive of the scale and the conformity in our judgments are the two forms of rigidity demanded for us (which reflect our causal situation in the world). Thus, the intersubjective sociability concerns a dimension of this rigidity that is necessary for us to be able to judge the world. According to Williams, the manners through which we are causally incorporated in the world provide the material for our normative practices, but we cannot reduce the normativity of our practices to the causal events beyond the very rigidity of the metal-bar that constitutes the “meter” of length.

Communitarianism³⁵ and classical behaviorism place human beings as a whole in the position of the learner: vulnerable and open to constant assessment. The intellectualist (cognitivist) perspective about the learning of the first natural

³² WILLIAMS, M. *Blind Obedience*. New York: Oxford University Press, 2014. p. 3.

³³ WITTGENSTEIN, L. *Philosophical Investigations*. New York: Blackwell, 2001, §104.

³⁴ WILLIAMS, M. *Normative Naturalism*. *International Journal of Philosophical Studies*. N° 18 (3), 2011. pp. 370-75.

³⁵ I refer specifically to Kripke’s skeptical solution presented in “*Wittgenstein on Rules and the Private Language*”.

language places human beings as a whole in the position of the competent user of the language. Anyone of these two exclusive manners of understanding our situation in the world ends up necessarily obliterating the crucial roles performed by what Wittgenstein called “background,”³⁶ and this leads us either to a strongly behaviorist image or to a strongly intellectualist image.

The process of natural language acquisition is, then, a process of enculturation within social practices, practices that implement the norms to “judge” reality and that allow us to perform movements within the language games. The process of learning does not occur within an epistemic context, but a context of calibration and adjustment to norms.

Well, let’s assume the child is a genius and itself invents a name for the sensation!—But then, of course, he couldn’t make himself understood when he used that word.—So does he understand the name without being able to explain its meaning to anyone?—But what does it mean to say that he has ‘named his pain’?—How has he done this naming of pain?! And whatever he did, what was its purpose?—When one says “He gave a name to his sensation” one forgets that a great deal of stage-setting in the language is presupposed if the mere act of naming is to make sense. And when we speak of someone’s having given a name to pain, what is presupposed is the existence of the grammar of the word “pain”; it shews the post where the new word is stationed. (PI §257)

What Wittgenstein is showing through his diagnosis of the intellectualist positions is exactly that “ostensive training” is a practice of transmitting abilities that necessarily precedes the intellectualized process of naming, that is, that it precedes the practice of “defining ostensively.” There are no doubts that natural language is an acquired activity, but before that happens, it is necessary to acquire the semantic and cognitive abilities to understand and follow the rules that are socially transmitted. Without developing such abilities, it is not possible to acquire a natural language.

Natural languages generally demand of its users very specific abilities. An empirical study conducted by Lera Boroditsky, Stanford University, explored the fact that native speakers of English and native speakers of Mandarin have different ways of describing time. Boroditsky showed that the abilities demanded by different natural languages are directly implied in the ways of describing reality. His experiment revealed that native speakers of English in general describe the passage of time horizontally (“good times ahead,” “long time ago” etc.). In contrast, native speakers of Mandarin usually describe the

³⁶ John Searle also highlights the importance of the term “background” in his “Consciousness and Language,” where he explicitly agrees with the position advocated here: “The fact that we consider certain acts to constitute the right application of a rule and others not; the fact that we consider certain acts to constitute a correct use of addition and others not, all of them are simply facts of our conduct, facts referring to the way we were reared and educated (...) What fixates the interpretation in the effective practice, in real life, is that which I have called ‘background’ elsewhere.” See: SEARLE, J. *Consciência e Linguagem*. São Paulo: Martins Fontes, 2010, p. 439.

passage of time vertically, using expressions such as “above” and “below” to speak of previous or subsequent months.³⁷

In another recent study, Boroditsky describes the case of a five-year-old girl in Pormpuraaw (a small aboriginal community in the west of Cape York, in northern Australia) who is able to immediately point to the north with accuracy and without hesitation while many renowned scientists cannot do so.³⁸ The “cognitive difference” is explained in terms of how a natural language molds thought, since the language of the natives of Pormpuraaw has no relative cardinal points, only absolute (north, south, east, west) and most expressions used necessarily reveal the location of the objects described. This results in the natural language of the aboriginal tribe more strongly demanding that its users have the cognitive ability of indicating the spatial location of their descriptions with accuracy.

The different abilities become more evident when we observe different ways of describing one same fact in different natural languages. To use an example based on Boroditsky’s research, suppose that S wants to convey that he has seen his aunt in 42nd street. If S is a native speaker of Mian in Papua New Guinea, then the verb used should reveal if the event has occurred in the present, in the recent past or in the distant past. If S is a native speaker of Bahasa Indonesia, then the verb used to perform the description will not reveal if the event has already happened or if it is still bound to happen. If S is a native speaker of Russian, in describing the situation he should reveal the gender. If S is a native speaker of Mandarin, he will have to specify if he is speaking about his maternal or paternal aunt (if she is related to him by blood or marriage), since there are many different words to all these types of relations. If S is a native speaker of Pirahã (a language spoken in Amazonia), he will not be able to say “42nd” since there are no words for exact numbers in Pirahã, only relative numbers (many, few, some, none).

3. FINAL REMARKS

Evidence presented in recent years to support Boroditsky’s thesis, in addition to challenging the thesis of the universality of language, also opens new perspectives about the origin of knowledge and the construction of reality. The fact that different natural languages demand specific abilities of their users also reinforces the thesis that these abilities are acquired through “ostensive training” in the enculturation process. These skills are acquired only due to our material properties (sensorial perception, neural plasticity etc.), which are susceptible to calibration. Therefore, material properties are necessary conditions to natural language acquisition, even if not sufficient. The ways of life, or even the shared social patterns, are the complementary elements to learning a natural language.

³⁷ BORODITSKY, L. *Does Language Shape Thought? Mandarin and English Speakers’ Conceptions of Time*. *Cognitive Psychology*, vol. 43, no. 1, pp. 1-22, Aug. 2001.

³⁸ BORODITSKY, L. *How Language Shape Thought? The Languages We Speak Affect our Perceptions of the World*. *Scientific American*, 2011.

The distinction between “ostensive training” and “ostensive definition” is crucial to explain the process of natural language acquisition, since it evidences that the abilities required by the natural language are transmitted through a non-semantic, or even pre-linguistic, element. If my considerations are correct, apriorism and innatism make the same mistake: they confuse “ostensive training,” as a non-intellectualized process of ability transmission, with the highly intellectualized process of ostensively naming or defining, a stage of language development that clearly can only work insofar as semantic and cognitive skills are already acquired and developed.

The notion of “ostensive training” is closely related to our most basic conventional patterns of activity. Ostensive training is naturally independent of problems of coordination, regular conformities and rational presuppositions. An important feature of this way of learning is exactly the fact that learning occurs by direct copy of conventional patterns of activity that, instead of rationally deliberated, are “blindly reproduced.” This means that cognitive abilities required by natural languages do not depend on any intrinsically superior mechanism to be acquired or developed. The naturalist perspective of explanation of the process of primitive language acquisition, advocated in this paper, has the important advantage of harmonizing causal determination of ostensive training with logical or normative determination of the enculturation process, which is not evidenced in other theoretical perspectives.

ON THE PROBABILITY OF THEISM

LUIS ROSA

1. The proposition expressed by ‘God exists’ (G), if it is true or false, is either necessarily true/false or not necessarily true/false. In other words, if G is capable of having a truth-value v , then it is either necessarily v or contingently v . By ‘God’ I mean a supernatural being, a powerful and immaterial mind that purportedly created the universe. Certainly there are other meanings that are attached to that term in certain contexts, but the arguments I will assess here only make sense under that interpretation. We can discuss arguments involving other meanings attached to the term ‘God’ on another occasion.²

Suppose that G has truth-value v . Now suppose that G is *necessarily* v (that is, G is *necessarily true* or G is *necessarily false*). That means that G cannot be v' , where $v \neq v'$ (e.g., v is the value *true* and v' is the value *false*). If so, then nothing can raise or lower the objective probability of G , $Ch(G)$. For example, if G is necessarily true, then not only $Ch(G) = 1$, but also $Ch(G | P) = 1$, for any proposition P . That alethic status, however, is possessed only by tautologies (e.g., *Fred is fred*), logical or mathematical axioms and theorems (e.g., $p \rightarrow p$, $(x + y) = (y + x)$), and so-called ‘analytical truths’ (e.g., that *Every bachelor is an unmarried man*). Similarly, the contrary alethic status of that status is reached only by *negations* of tautologies, axioms/theorems and analytic truths.

But G is no tautology/negation of a tautology. And it is neither a mathematical or logical axiom/theorem, nor a mathematical or logical contradiction. Still, one might be tempted to suggest that G could be an analytic truth. Consider typical examples of analytical truths. No one who is a competent user of the terms ‘bachelor’, ‘unmarried’, ‘man’, etc. seriously asks whether it is true that *Every bachelor is an unmarried man*. That proposition is sufficiently obvious to every such competent speaker. But G is by no means obvious (the fact that there is so much controversy around it testifies that), and we seriously ask about its truth.

¹ Originally published in *Dissertation's Journal of Philosophy, Naturalism Dossier 2015*.

² I often feel like it becomes completely pointless to talk about the existence of God under certain interpretations of the term ‘God’. For example, some people say ‘God is everything there is’. It is true that God exists in *that* sense, of course. But this is not what we want to know when we ask whether God exists or not—we are not asking whether everything there is exists. (Similar observations apply to other cases involving supernatural terms, e.g., when we want to know whether spirits or souls exist but some people say that the spirit is a *physical* energy, or something like that).

Of course, assuming that there is a substantial distinction between analytic and synthetic truths, it might be the case that some analytic truths are not obvious at all. Maybe G is one of them. The proponent of that line of thought, however, would have to face two challenges. First, it has been pointed out that the relevant distinction does not really cut human thought and language at its joints.³ Second, purported analytic truths that are somehow discovered or fleshed out are not really surprising *after* they have been discovered or fleshed out. Even when there is disagreement between conceptual analysts, they recognize each other's conceptual analyses as being very close to the establishment of necessary and sufficient conditions envisioned in their analytical endeavours. But if a conceptual analyst were to present us something like: x is God if and only if (i) x exists, (ii) such-and-such, (iii) so-and-so, etc., although we would probably be surprised, we would not accept it as a suitable analysis of 'God'.

It follows, then, that whatever truth-value G has, it does not necessarily have that value. Let q be the value such that $Ch(G) = q$ (that is, the objective probability of G is q). If what we said about the alethic status of G is right, then it is possible for there to be certain facts P such that $Ch(G | P) < q$, and it is also possible for there to be certain facts P' such that $Ch(G | P') > q$.

Given certain assumptions (involving both the syntax and the semantics of the formulas that constitute the arguments of probability functions), we can say the same thing about probability functions under other interpretations. In particular, we could say the same thing about an epistemic probability function Pr . Here, of course, we talk about *evidence* instead of *facts* (without implying that no evidence is factive, of course): given $Pr(G) = q$, there may be some piece of evidence E such that $Pr(G | E) < q$, as well as some further piece of evidence E' such that $Pr(G | E') > q$. (From now on, whenever I use the term 'probability' I mean *epistemic* probability).⁴

So, just as we can say that the hypothesis that God exists is neither necessarily true nor necessarily false, so we can say that it is neither certainly true nor certainly false conditional on our evidence. This is part of what is involved in the claim that G is an *empirical hypothesis*, or that G is a hypothesis about how the world is constituted.

2. If that is right, then we can build inductive arguments pro or con the existence of God. Maybe we can find pieces of evidence that *raise* the probability of G ; maybe we can find pieces of evidence that *lower* the probability of G ; maybe we can find *both* types of evidence, and these different pieces of evidence will 'fight' for evidential support. But it may also be the case that there is no evidence available to us that has any bearing on the probability of G .

³ E.g., see Quine's classic (1951) and also Williamson (2007, Ch. 3–4). A good starting point for a problematization toward that distinction is to ask: What does it mean to say that a sentence or proposition is true 'solely in virtue of meaning', and not in virtue of how the world is?

⁴ Probability functions are functions that obey to the usual Kolmogorov axioms—see Hájek (2002).

Theists—those who claim that *God exists*—purportedly have good undefeated evidence that gives overall support to G . For any proposition Φ and body of evidence Γ , let r be the value such that, if $Pr(\Phi | \Gamma) \geq r$, then a subject whose total evidence is Γ is entitled to believe that Φ .⁵ Theists would need to possess, then, some body of evidence E such that $Pr(G | E) \geq r$, and they would need to possess no further piece of evidence E' such that $Pr(G | E \wedge E') < r$ (presumably, the evidence E mentioned here is evidence that theists can share with their fellow human beings).

Atheists—those who claim that *God does not exist*—purportedly have good undefeated evidence that gives overall support to $\sim G$. They would need to possess, then, some body of evidence F such that $Pr(\sim G | F) \geq r$, and they would need to possess no further piece of evidence F' such that $Pr(\sim G | F \wedge F') < r$ (again, the evidence mentioned here is presumably evidence that atheists can share with their fellow human beings).

While there is just one way to go for the theist and one way to go for the atheist, the *agnostic*—or *skeptic* if you like—has more than one way to go. The agnostic refrains from claiming either that *God exists* or that *God does not exist*—he/she suspends judgment about G . In order for suspension of judgment to be the rational attitude for one to take toward a certain proposition, one's evidence need to be sufficiently neutral as to whether that proposition is true. There are two ways in which that might be the case.

First, one may be in a situation of *evidential poverty*—a situation, that is, in which one possesses no evidence pro or con the truth of a certain proposition. In the case at hand, agnostics may possess no body of evidence D such that either $Pr(G | D) \geq r$ or $Pr(\sim G | D) \geq r$.

Second, one may be in a situation of *evidential symmetry*—a situation, that is, in which one possesses some evidence pro and some evidence con the truth of a certain proposition—but each counterbalances the other. Such a situation would be modeled as follows. Suppose a subject S possesses some evidence E such that $Pr(G | E) \geq r$. If S 's total evidence were E , then S would be epistemically entitled to believe that G . As it happens, though, S has some further body of evidence F such that $Pr(\sim G | F) \geq r$, which is the same as to say that $Pr(G | F) \leq 1 - r$.

Now, notice that just as the threshold on $Pr(\Phi | \Gamma)$ for epistemic entitlement to *believe* that Φ is r (assuming that Γ is the total evidence), so the threshold on $Pr(\Phi | \Gamma)$ for epistemic entitlement to *disbelieve* that Φ , or to believe that $\sim\Phi$, is $1 - r$ (assuming again that Γ is the total evidence), for any proposition Φ and body of evidence Γ . So for any value s in the open interval $(1 - r, r)$, if $Pr(\Phi | \Gamma) = s$ and Γ is S 's total evidence, then S is entitled to suspend judgment about Φ . Just as there are thresholds on epistemic probabilities for epistemic entitlement to believe/disbelieve a proposition, so there are thresholds on epistemic

⁵ In order for that consequent to hold, we should also assume that the subject is *capable* of forming a belief toward Φ on the basis of evidence Γ (no one is entitled to believe something that one can only believe in a non-competent way). But this is not particularly important now.

probabilities for entitlement to doubt or to suspend judgment about a proposition.

Now apply that result to the scenario we have built above (evidential symmetry). We assumed that S has some evidence E in support of G and also some evidence F in support of $\sim G$. Assuming that E and F are all the relevant evidence that S has concerning the truth of G , if $1 - r < Pr(G | E \wedge F) < r$, then S is entitled to suspend judgment about the proposition that *God exists*.

So agnostics have two argumentative resources at their disposal: either we are in a situation of evidential poverty with respect to the hypothesis that *God exists*, or we are in a situation of evidential symmetry with respect to that hypothesis.

3. Some philosophers have explicitly offered probabilistic arguments pro the existence of God, using a bayesian framework. Most notably, Richard Swinburne (2004) makes use of traditional bayesian criteria for establishing the confirmation of hypotheses and the explanation of observable evidence, and argues that we have overall good evidence for believing that *God exists*. Among these criteria, two of them will be particularly relevant for the points I want to make.

First, how much support a certain body of evidence E confers upon a certain hypothesis H is dependent on the *prior probability* of H . This becomes clear as we look at Bayes' Theorem:

$$Pr(H | E) = Pr(H) \times Pr(E | H) \div Pr(E).$$

As we increase the value $Pr(H)$ and maintain the rest, so the value $Pr(H | E)$ increases. As Swinburne himself points out (2004: 53), the prior probability of a hypothesis—or its probability *before* we have made some relevant observations E —is determined by its *fit* with our background knowledge B , its *simplicity* and its *scope* of application. The background knowledge B is our general knowledge of how the world works, and it is supposed to figure as follows:

$$Pr(H | E \wedge B) = Pr(H | B) \times Pr(E | H \wedge B) \div Pr(E | B).$$

Now, a hypothesis H is said to *fit* the background knowledge B when it postulates similar types of entities, properties, relations and laws as the ones present in B . Similarly, the simplicity of H is a matter of its not postulating more entities, types of entities, properties, relations and laws than the ones present in our background knowledge B . So H is supposed to be as concise and economical as possible. Finally, the scope of a hypothesis is the range of objects about which it has implications. We will get back to these points in a moment.

Second, how much support a certain body of evidence E confers upon a certain hypothesis H is also determined by the explanatory power that H has with respect to E . (This again becomes evident when we look at Baye's Theorem: as we increase the value $Pr(E | H \wedge B)$ and maintain the rest, so the value $Pr(H | E \wedge B)$ increases). Explanatory power is a matter of how likely the

hypothesis makes the observable evidence. If H entails E , for example, it cannot get better than that: the hypothesis makes sure that the evidence occurs.

So in the case at hand, where we want to determine the probability that *God exists* given our available evidence, we need to answer two crucial questions: (i) What is the prior probability of G ? (ii) How much explanatory power G has with respect to our shared observable evidence? When we address (i), we must check how well does G fit with our common background knowledge, how simple G is and what the scope of G is. When we address (ii), we must check how likely is our shared observable evidence conditional on the existence of God. Let us turn to those questions now.

4. Let us consider, first, how well does G fit with our background knowledge B . In order to do that, we must at least roughly establish what our background knowledge about the world is supposed to be. It will contain, among other things: causal information (e.g., normally fire causes smoke),⁶ temporal information (e.g., my parents were born before me), spatial information (e.g., the Earth and the Moon occupy and move through different regions of the solar system), statistical data (e.g., all observed ravens are black), psychological data (e.g., about how things feel and appear to us), conceptual relations (e.g., all humans are mortal), logical and mathematical principles (e.g., if P and $P \rightarrow Q$ then Q , if $x > y$ and $y > z$ then $x > z$).

That much will suffice for our present purposes.⁷ The most important point is that B is supposed to contain only knowledge that is *maximally common* between us. The more controversial a certain claim is, the more unqualified it is to be inserted in B .

Now, God is supposed to be an immaterial being, a conscious mind with no body or matter. God would have, for example, beliefs and intentions, as God is supposed to have knowledge of all knowable facts and also to be benevolent. But God would not need to have a *nervous system*, for example.

In our background knowledge, however, there is not a single example of a being with consciousness and intentionality *but no* such material structure. Statistically, *every* mind we purportedly know to exist has a physical medium (a

⁶ The inclusion of causal information in B is independent of the fact that we may have a deflationary interpretation about causal claims (e.g., to the effect that purported causal relations are nothing more than associations or correlations between events/types of events). That observation generalizes to the other types of information that are included in B .

⁷ This is what we call 'common background knowledge about the world'. Swinburne suggests that we start the probabilistic investigation about G with a background knowledge devoid of any substantial content, with nothing but tautologies in it, and then start enriching it. But this has at most heuristic value for Swinburne's purposes, and at the end of the day it does not make a probabilistic difference if we start computing the 'pure' prior probability of G and update it on substantial background knowledge or if we already start computing the prior probability of G with that substantial information in B .

medium by means of which that very mind manifests itself).⁸ No cognitive agent known to us lacks a brain. And no cognitive agent known to us lacks representations, or vehicles of information. How is God supposed to know things if God does not represent the world as being a certain way? And, if God *does* process information and handles representational items, how does he do that without a brain or something analogue to a brain?

On the face of that statistical data (all minds known to us have bodies and all of them process information), which is part of our background knowledge, it turns out that $Pr(G | B) = p$ must be quite low. If we had started with a pure prior $Pr(G) = 0.5$, for example, we would not only have $Pr(G | B) < Pr(G)$, but also quite a substantial difference $0.5 - p$. Assuming again that r is our upward threshold of evidential support that entitles one to believe a proposition, and that $1 - r$ is our downward threshold of evidential support that entitles one to disbelieve a proposition, it could well be the case that $p \leq 1 - r$, in which case we would be entitled to disbelieve G . But since we are making pretty coarse grained estimations here, it could also be that $1 - r < p < r$, in which case we would be entitled to suspend judgment about G . Either way, we would not be entitled to believe that G conditional on B . As far as B goes, then, theism would not be the rational position.

That would not be the *only* information in B that lowers the prior probability of G . The situation gets worse, at least if we assume that God is supposed to be the creator of the universe. Remember that the simplicity of G is supposed to have a direct influence on the prior $Pr(G | B)$, and that the simplicity of a hypothesis is a matter of its not postulating more entities, types of entities, properties, relations and laws than the ones present in B . But if God is supposed to be the creator of the universe, then he must have established some causal relation with the universe. But God is immaterial and the universe is material. So G postulates a type of causal relation that differs from the one present in our causal claims in B .

Consider the causal facts that we ordinarily take ourselves to know. We purportedly know that my throwing the ball into the window may cause it to break, that the wind may lower the temperature of a certain place, that certain drugs may bring about health improvement/decay, etc. In these examples, what we have is physical causation: one physical event (or a whole set of physical events conjoined) causes a further physical event. In none of these cases we have a spiritual substance interacting causally with a material one.

You may think that mental causation fits the bill, though. You may think, that is, that our minds are immaterial substances that interact causally with our bodies, e.g. when desires purportedly cause bodily movements—in which case you are a *substance dualist*. That there is causation between a spiritual substance and a physical one in our cognition, however, is not itself part of our common knowledge (as *substance dualism* surely is not part of our common knowledge).

⁸ This is not to say that there is no immaterial substance. So far, we are just pointing out that there is no mind unaccompanied by a body in our background knowledge—and this is all that is needed for making our point here.

For all we know, mental causation might be physical causation as well. It is not the case that everybody or almost everybody knows that spirits cause physical events—this is quite a controversial matter.

So claims about immaterial substances maintaining causal relations with physical ones is not uncontroversial enough to be inserted in B . As long as G postulates a new type of causal relation—an unknown type of relation—the prior $Pr(G | B)$ will be even lower, for in that case G does not fit B .

5. In the previous section we made the case that the prior probability of G is quite low. In the present one, we will make the case that G does not fare better when it comes to explanatory power. The problem concerning the explanatory power of theism is also related to problems with substance dualism.

The so-called mind-body problem was one the main reasons (maybe *the* main reason) for the demise of substance dualism throughout 20th century philosophy. What is the relationship between mind and body? If the mind is an immaterial substance and the body is a material one, how could they interact causally, if at all? Nowadays, we have versions of dualism (e.g. epiphenomenalism and parallelism) that try to avoid commitment to the idea that there is causal interaction between mind and body.

The thesis that God is the creator of the universe has ontological commitments in common with substance-dualism. Here we have the God-universe problem: what is the relationship between God and the universe? If God is immaterial and the universe is material, how could they interact causally, if at all? How could God produce movement and heat, for example? Maybe God could use a stick or some sort of dummy member as a tool. That way, we could have a bridge between God and the universe. But where did the stick/dummy member come from? And how could God interact with something even before that thing already exists? So it would seem that the probability that there is a material world conditional on the existence of God is very low indeed. Indeed, if it is *impossible* for the spiritual realm to interact causally with the physical realm, then that probability is 0.

But if theism is supposed to have at least some explanatory power with respect to the observable evidence (the universe as we perceive it), then it must be because God created the universe as it is or, maybe, because God interacts with it, influencing the course of history. Could God indeed be the creator of the universe that is observable to us? Could God indeed interact causally with the universe?

Given again that we know of no causal relationship between spiritual and material substances, both things are unlikely conditional on B . Let C be the proposition that *God created the universe*, or *God interacts with the universe* (it does not really make a difference for the present purposes which one you choose). Then, for reasons similar to the ones presented in the previous section, we would have a low $Pr(C | B) = p$. Indeed, we only refrain from assigning $Pr(C | B) = 0$ because there is a bare possibility of causal interaction between spiritual and physical substances that is purportedly left open by B . Now, this is not the relevant result we want to draw *per se*—the issue here is the explanatory power

of theism. The relevant conclusion to be drawn is the following. Let E be a conjunction of evidential claims that can be summed up as *The universe exists, and it appears to us to be a certain way w* (where w describes the natural properties/relations we seem to observe in nature). Now, E is actually entailed by B , so it would be fruitless to test the explanatory power of theism by computing the likelihood $Pr(E | G \wedge B)$. For here B ‘trumps’ the relevance of G to the truth of E .

The same applies of course to other hypotheses, e.g., cosmological theories. We have in our background knowledge propositions that entail that there is a universe, that the universe has such-and-such properties, etc. And what we expect from an empirical hypothesis is that it is able to *explain* the observable evidence. Since in the bayesian framework this is a matter of the likelihood of the hypothesis, $Pr(E | H)$ or $Pr(E | H \wedge B)$ depending on the case, we can test the explanatory power of an empirical hypothesis by stripping out information from B that entails E . In this way we can maintain a smaller set of propositions that are part of B and generate a thinner, alternative background set B^- . Importantly, we can keep generalizations, conceptual relations and logical or mathematical claims in B^- . So, for example, the claim that *All the causal relations that apparently manifested themselves so far are physical relations* will be maintained, and so will *x is God if and only if x is immaterial*.

Given that much, if it is right to say that $Pr(C | B)$ is low, then we can also conclude that $Pr(E | G \wedge B^-)$ is low: if it is quite unlikely that God, which is an immaterial being, creates or interacts with the universe, then it is quite unlikely that there is a universe like the one we seem to observe conditional only on the existence of God and the information that ‘survived’ our selection to generate B^- (what we stripped out from B , remember, is solely a set of propositions that entail E). It would follow, then, that the hypothesis that *God exists* is not able to play any explanatory role with respect to E for us.

6. CONCLUSION

So far, however, we can only conclude that *probably* $Pr(G | E \wedge B) < r$, and so that we are *probably* entitled to *disbelieve* G , that is, to believe that $\sim G$.⁹ But I want to be cautious and I do not want to prematurely jump to conclusions. This paper will have a sequence. Until then, I will follow the good old skeptical advice of suspending judgment on such complicated matters.

⁹ In this case, ‘probably’ means ‘as far as I can see’ or *somesuch*.

NATURALISM AND REALISM

MARIO DE CARO

In the last decades, many version of liberal or non-reductive naturalism have been developed. Some of these forms – often Kantian or Wittgensteinian in spirit – defend a quietist attitude, by not taking a stance regarding the issues normally discussed in the context of the discussions on realism and antirealism. Other forms of liberal naturalism, however, take an explicit realist attitude. In this article, I will discuss the latter conceptions, arguing that they are the most promising ones.

1. REALISMS AND ANTIREALISMS.

The problem of realism should not be stated in an “all or nothing” form. In fact, no philosopher has ever been a complete realist or a complete antirealist. Take for example Alexius Meinong, possibly the most fervent of all realists: even he believed that a round square could not be real. Or, on the other hand, take Berkeley, a champion of antirealism – but only in regard to matter, since, as to the mind (especially the divine mind), he was an arch-realist. All serious philosophers are located in the interval between a hypothetical integral realism and an equally hypothetical integral antirealism. All attempted solutions of the problem of realism are therefore matters of degree. The problem is to determine what is the right dose of realism to be taken – and this of course is not an easy task.

The complication of the issue of realism is also due to other reasons. The first is that one can be committed to realism in many different forms. Fortunately, however, the forms of realism can be grouped in three main families:

(i) *Ontological realism*. The views of this family are based on the thesis that things of certain sorts, be they concrete or abstract (such as disembodied minds, numbers, witches, electrons, and social facts), be they properties (redness, goodness, electric charges, or free will) or events (the Big Bang or transubstantiation) are real. Theories of this family can also claim the existence of the external world as a whole or specify in what sense time exists (are the past and the future real or is the present the only real time?). Moreover, when discussing realism from the ontological point of view, one can ask two distinct questions. One can ask if a certain thing really exists or, granting that it does, one can ask whether it exists independently of the minds that think about it. For instance, with regard to atoms it is the first question that standardly arises: do atoms truly exist or are they only heuristically useful fictions? Conversely, on

the subject of colors it is the second question that arises: do colors exist independently of us, out there in the world, or do they exist only – as according to Galileo, Locke and many contemporary philosophers –to the extent that a mind projects them onto the world?¹

(ii) *Semantic realism*. According to the views of this family, the meaning of a sentence is given by the conditions under which it is true, and this requires the reality of the entities to which the terms of that sentence refer. Semantic antirealism claims that the meaning of a sentence is instead given by the conditions under which speakers are justified in asserting it.

(iii) *Epistemological realism*. According to the views of this family, there may be facts that are, in principle, unknowable (whereas epistemological antirealists deny this possibility).²

Another reason why this discussion is complicated is that, while it is more or less clear what realism (in the three senses described above) amounts to, the term “antirealism” labels a much more disparate collection of views. Among the different forms of antirealism one should include nominalism, idealism, phenomenism, conventionalism, relativism, radical empiricism, and skepticism, almost all of which are brought forward in the various ontological, epistemological, and semantic discussions to challenge realism in different and often insidious ways.³

As said above, in all philosophical views there are elements of realism and antirealism. This remark is useful in order to analyze the two most general versions of realism in vogue today, namely: (i) *common sense realism*, the view that acknowledges as real only the things we can experience (directly, with the senses, or indirectly, with the instruments that extend them), and (ii) *scientific realism*, the view that claims that the world contains only the entities and events (both observable and unobservable) that are in principle describable by science. When taken in their pure forms these are alternative realist conceptions because each of them is drastically antirealist in the field in which the other assumes a firmly realist attitude. As we will see, however, these views are restatements of two traditional (and unsatisfactory forms of realism); consequently they do not deserve the title ‘New Realism’, as instead does another form of realism that I will discuss in the last part of this article.

In order to compare the virtues and vices of these two conceptions, it is useful to look at how they respectively answer a problem that has afflicted

¹ When discussing the independent existence of an entity, the question is not to be posed in genetic or causal terms. Obviously the table in front of me exists because someone has built it; however, once it is built, the table exists independently of its builder.

² The definitions of epistemological and semantic realism given above are inspired by Michael Dummett’s (1978) seminal work. A different definition of epistemological realism, common in the debate in philosophy of science, is based on the idea that the theoretical claims that concern unobservable entities can constitute knowledge of the world (see Chakravarty 2007, 9).

³ See Chakravarty (2007, 9-13) for a taxonomy of the different forms of realism and antirealism that – if specifically considered within the realism/antirealism debate regarding the natural sciences – can actually be of a general interest.

philosophy since the age of the scientific revolution – a problem that Kant presented in the most illustrious way in the third antinomy of the *Critique of Pure Reason*. As Kant put it, the problem was that we cannot help believing in two contradictory views of the world: one in which we are free and responsible agents and another in which everything is ruled by the ubiquitous laws of nature, which in his opinion do not leave any room for freedom and responsibility. However, Kant’s own ‘transcendental’ solution to this problem has been found convincing by very few, and so the discussion has continued until today. Here is a reformulation of the problem given by Searle (2007, 4-5):

How can we square [the] conception of ourselves as mindful, meaning-creating, free, rational, etc., agents with a universe that consists entirely of mindless, meaningless, unfree, nonrational, brute physical particles?

Both the common sense and the scientific views of the world have been used in the attempt to solve this problem, by opting for the unilateral realistic options that they respectively embody. According to the prototypical common-sense realist, the real world consists in just what the ordinary view of the world assumes: colors, sounds, smells, and perceptible physical bodies; values, too, may be held to have an objective reality along with intentional, conscious, and moral properties, which all are assumed to be irreducible to scientific properties. Vice versa, common-sense realism sees the unobservable entities postulated by scientific theories as nothing more than fictions – useful for producing explanations and predictions, but with no ontological legitimacy. The scientific realist, on the contrary, assumes as real only the entities admitted by the natural sciences, while all other alleged entities either are reducible to those or are mere fictions.

In my opinion these views, with their respective versions of unilateralism, represent the two main forms of “Old Realism”. Let us consider them briefly, before turning to the form that a satisfying New Realism should assume in order to overcome the shortcomings of these older realist conceptions.

2. COMMON SENSE REALISM AND SCIENCE.

Common sense realism has a long and authoritative history. Among its most influential defenders are Aristotle and most of the Aristotelian tradition, Reid, William James, Pierre Duhem, G. E. Moore, P.F. Strawson, and J.L. Austin and, in some relevant aspects of their thinking, Edmund Husserl, Henri Bergson and Ludwig Wittgenstein. Of course, there are differences, and even important ones, between the views defended by these philosophers (Husserl’s *Lebenswelt* is not the same as Moore’s commonsensical world or James’s pragmatist reality). However, all these stances can be seen as important predecessors of contemporary common sense realism, a view that is based on two main thesis.

The first thesis is that perception tends to give us access to the external world as it really is. In this perspective, apart from special cases – such as optical illusions or the situations in which perceptual conditions are not optimal – observable objects really do have the properties that, on the basis of perception, we tend to attribute to them; and this is true of both primary properties such as

size and shape and secondary ones such as color and odor. The second thesis of common sense realism is that the middle-size objects we perceive have properties that are not identical to whatever microphysical properties constitute them. Consider the case of a table in front of me: its functional characteristics and aesthetic value are neither identical nor be reducible to its constitutive microphysical properties. Consequently, no description of the table that only mentions its physical properties could account, even in principle, for its functional or aesthetic characteristics.

Precisely because common sense realists attach so much importance to perception – by assuming that it is a reliable guide to the nature of the objects and properties of the external world – they standardly assume an anti-realistic attitude towards the unobservable entities posited by natural science (such as electrons, radiation or black holes). This is understandable, since from the point of view of common sense these alleged entities exhibit incomprehensible behavior.

In his *The Crisis of European Sciences and Transcendental Phenomenology* (1936), Edmund Husserl gave a genealogical account of the antirealist view of the common-sense world produced by modern science. In his opinion, Galileo was responsible for

the surreptitious substitution of the mathematically substructured world of idealities for the only real world, the one that is actually given through perception, that is ever experienced and experienceable – our everyday life-world. This substitution was promptly passed on to his successors, the physicists of all succeeding centuries. (Husserl 1936, 48).

According to Husserl, the only real world is the “life world”, the world of human experience – a world in which value and meaning are real, and the so-called secondary qualities do belong to the external objects in which we would commonsensically locate them (see Moran 2008). This is the “forgotten meaning-fundament of natural science” (Husserl 1936, 48), since scientific concepts are mere idealizations with practical purposes, such as measurement and prediction, but do not refer to any unobservable reality. On this perspective, science can at best be interpreted instrumentally, that is, in antirealist terms.⁴

Husserl was very explicit in his association of common sense realism with a strong antirealist attitude toward science. Less explicitly, the philosophers of the hermeneutic tradition (including Heidegger, Gadamer, Ricoeur, and Derrida) also took the framework of common sense as the starting point of their analyses, trying to re-evaluate its underlying assumptions. In this regard, the hermeneutic philosopher Claudio Ciancio convincingly writes that,

There is no doubt that the topic of common sense plays an important role for the thinkers of a hermeneutic orientation, or at least for some of them, although in their writings there is a fairly limited mention of it. What attracts hermeneutics to the issue of common sense is its antirationalistic and antiscientific orientation

⁴ Useful discussions on this issue can be found in Willard (2002) and Wiltsche (2012).

aimed at developing a notion of the concrete universal. (Ciancio 2004, 153; also Bunge 2006, ch. 3).

Therefore, in the case of at least some continental philosophers, the dismissal of scientific realism was the product of an ideologically antiscientific attitude. This, however, certainly cannot be said of Bas van Fraassen, one of the leading contemporary philosophers of science, whose “Constructive Empiricism” is another influential example of the conjugation of common sense realism and antirealism with respect to science. According to van Fraassen, a scientific theory that appeals to unobservable entities and produces sufficiently accurate explanations and predictions of observable evidence should not be considered a true description of the world, even if its “empirical adequacy” makes it an extremely useful heuristic tool.

It is interesting to note that, in contrast to many other versions of antirealism in philosophy of science, van Fraassen’s instrumentalism does not derive from the traditional empiricist view according to which our knowledge of the external world coincides with the knowledge of our sense data. According to van Fraassen, scientific antirealism is rather connected to traditional common sense realism and its tenet that we have direct knowledge of the observable world: “Constructive empiricism is set squarely within a common sense realism that was foreign to much of the empiricist tradition.” In this light “[t]he common basis I assume is language in which reference is unproblematic to trees and mountains, people and books.” (van Fraassen 2003, 479).⁵

Therefore van Fraassen’s view represents one of the most consistent expressions of common sense realism, since it limits the scope of the knowable to what is directly observable and, accordingly, assumes an antirealist position with regard to science precisely insofar as it refers to unobservable entities.

Husserl and van Fraassen tried to show how to account for what science tells us about the world once one has given epistemological and ontological primacy to common sense. However, their attempts are not convincing precisely because of their unilateralism, that is, their commitment to an antirealistic view of science. As a matter of fact, today there are excellent reasons for assuming a realist attitude toward science. The first is the famous “no-miracles argument,” originally proposed by Hilary Putnam (1975, 73). This argument is based on the idea that the only way of explicating the great explanatory and predictive success of the best theories of modern science is to acknowledge that these theories tell us the truth (or a good approximation to the truth) regarding the natural world. If one instead assumes an antirealist perspective, the fact that science functions so well – as it offers comprehensive explanations and extremely precise predictions – becomes an inexplicable mystery or, as Putnam puts it, a miracle. Therefore, it is rational to take our best scientific theories as true and, consequently, to welcome the entities these theories presuppose as real, even when they are unobservable.

⁵ A similar view was expressed by another Catholic thinker, Pierre Duhem, who is reported to have called himself “the unceasing apostle of common sense” (Martin 1991, 89).

Unsurprisingly, antirealists have tried to attack the miracle argument in various ways, but these objections can, in my view, be rebutted.⁶ At any rate, even if one is not happy with the no-miracles argument, there are other excellent reasons for preferring realism to antirealism in regard to science. First, it has been convincingly argued that, even in its most sophisticated versions such as that of van Fraassen, scientific antirealism leaves one vulnerable to discredited views such as solipsism and phenomenalism (Forrest 1994). Second, several appealing new versions of realism have been offered in recent years. Arguably, the most relevant is *structural realism*, the view that our best theories do not describe the intrinsic nature of the unobservable phenomena to which they refer, but rather their structure, i.e., the relations these phenomena enter into (Worrall 1989; Ladyman 1998, 2014). This is a vague definition, of course, and it has been refined in different ways, the most important of which define the ‘epistemic’ and the ‘ontic’ forms of structural realism. The former view claims that we can only know the structural (i.e., relational) aspects of the unobservable physical reality but not the intrinsic nature of the objects that are in relation with each other; the latter claims, more radically, that there are no unobservable objects but only structural features (Chakravartty 2004). Structural realism, in both its forms, is widely discussed today and has many advocates.⁷ Here, however, it is sufficient to notice that scientific realism is more alive than ever, and it is reasonable to believe that the burden of proof is on those who pretend to ignore or even deny it.

Defending the common sense view of the world is a very commendable philosophical goal, but doing that by sacrificing scientific realism is too high a price to pay. Nowadays, a satisfying New Realism should give science the ontological credit it deserves.

3. SCIENTIFIC REALISM AND COMMON SENSE.

As I said, today there are excellent reasons to embrace scientific realism; it is no surprise, then, that a vast number of philosophers (perhaps even a majority) have done so. Moreover, many of these philosophers take scientific realism as the main pillar of a very common metaphysical view, ‘strict naturalism’ or ‘scientific naturalism’, which in my opinion is one the main contemporary incarnations of Old Realism.⁸

⁶ See Chakravartty (2007). The main charges against the miracle argument are that it is based on an inference to the best explanation (which allegedly is either doubtful in itself or not applicable at this metaexplanatory level); that different theories can account for the same sets empirical data; and that the history of science shows that false theories (such Ptolemaic astronomy) can offer excellent predictions. Against these arguments, see Musgrave (1988), Psillos (2011), Putnam (2012), De Caro (2011), Alai (2014).

⁷ On epistemic structural realism, see Worrall (1989) and Morganti (2004); on ontic structural realism, see Psillos (2001, 2006), French (2006), and Chakravartty (2007a, ch. 3).

⁸ See De Caro and Macarthur (2004a; 2004b; 2010). In those writings, the preferred name for this view was ‘scientific naturalism’, but to avoid confusion with ‘scientific realism’ I will instead use the synonymous label ‘strict naturalism’.

Strict naturalism rests on four main tenets:

(i) An *ontological tenet*, according to which reality consists of nothing more than the entities to which the successful explanations of the natural sciences commit us;

(ii) An *epistemological tenet*, according to which scientific inquiry is our only genuine source of knowledge; all other alleged forms of knowledge (e.g., ordinary perception, *a priori* knowledge and introspection) are either reducible in principle to scientific knowledge or illegitimate;

(iii) A *semantic tenet* according to which no linguistic term can refer to entities that are both not acceptable to the natural sciences and irreducible to those which are;

(iv) A *metaphilosophical tenet*, according to which philosophy must be continuous with science as to its contents, methods, and purposes.⁹

Not only does this view presuppose the correctness of scientific realism (what the best scientific theories tell us about the world has to be taken for true, including the existence of the non-observable entities they presuppose), but it also assumes that the reality described by science is the only reality there is. Unsurprisingly, then, the main problem that strict naturalism has to face is the so-called “location problem” (Jackson 1998, 1-5) or “placement problem” (Price 2004). This problem concerns those features of the common-sense view that, at least *prima facie*, do not fit into the scientific view of the world (features such as free will, moral properties, normativity, consciousness, and other elusive phenomena such as financial debts or collective intentionality).¹⁰ In the perspective of strict naturalism, either these features are reducible (perhaps after undergoing a ‘revisionary’ treatment) to scientifically acceptable features of the world or they are mere fictions, and should be treated as such – that is, either taken as helpful but illusory beliefs or should be abandoned (as we did with the idea that the Earth does not move), depending on whether they play a useful, and perhaps indispensable, social role or not.¹¹

⁹ In contrast to De Caro and Voltolini (2010), this list of the commitments of scientific naturalism also includes a semantic tenet, which I now think is useful to distinguish from the others.

¹⁰ Smith (2003) points out that entities belonging to the domain of finance such as electronic money and debts are one of the most difficult challenges for the strict naturalist. Another complex case is that of mathematical entities (which many mathematicians, it should be remembered, take for real). The relation between mathematical realism, on the one side, and scientific and commonsense realism, on the other side, is a very complicated issue (see Maddy 1990 and Burgess 2008). At any rate, assuming that mathematical entities are real abstract entities (as even Quine does), means moving away from orthodox scientific naturalism, since this view only accepts the existence of entities located in the spatio-temporal world studied by the sciences of nature (see Field 1980, for a fictionalist treatment of mathematical entities in line with a rigorous version of scientific naturalism).

¹¹ An example in this sense is offered by the discussion between the authors that claim that our belief in free will is illusory. Some (such as Smilansky 2002 and Strawson 2010) argue that it would be undesirable and practically impossible for us to abandon that belief, others (such as Honderich 2002 and Pereboom 2014) affirm the opposite.

The main sources of inspiration for the rise of strict naturalism have arguably been the philosophies of Quine and Sellars. Today not every strict naturalist is a physicalist,¹² but many are, and their debt to Quine is obvious. In a physicalist spirit, he famously claimed that philosophy should be methodologically continuous with science (“Normative epistemology is a branch of engineering,” Quine wrote [1985, 664], in order to stress the point that there is not such thing as an irreducibly normative epistemology). Both ontologically and epistemologically, microphysics should have the last word:

Nothing happens in the world, not the flutter of the eyelid, not the flicker of a thought, without some redistribution of microphysical states ... If the physicist suspected that there was any event that did not consist in the redistribution of the elementary states allowed for in his physical theory, he would seek a way of supplementing his theory. Full coverage in this sense is the very business of physics, and only of physics. (Quine 1981, 98).

Quine still has a significant influence on many strict naturalists also because of his famous criterion of ontological commitment, according to which we should only accept the existence of the entities that have to exist for our best physical theories to be true.¹³

What is even more interesting here, however, is the role that Wilfrid Sellars played in the development of contemporary naturalism. Sellars’s accounts of the relations between the “manifest image” (the world as it is understood by common sense) and the “scientific image” (the world as it is understood by natural science) is diametrically opposed to that offered by Husserl in the *Crisis*. And this is no coincidence since, when he was a student at Buffalo, Sellars was deeply influenced by Marvin Farber, an heterodox phenomenologist who had studied with Husserl:

Marvin Farber... introduced me to Husserl. His combination of utter respect for the structure of Husserl's thought with the equally firm conviction that this structure could be given a naturalistic interpretation was undoubtedly a key influence on my own subsequent philosophical strategy. (Sellars 1975, 283).¹⁴

Like Husserl, Sellars tried to understand the relation between the ways in which common sense and science respectively conceive of the world; and, like Husserl, he aimed at finding a unified view (a “stereoscopic vision”, he called it). In this regard, he wrote that these two images are “pictures of essentially the same order of complexity, each of which purports to be a complete picture of man-in-the-

¹² A scientific naturalist can actually believe at the same time that the natural sciences have ontological and epistemological primacy and that some natural sciences (i.e., biology) are in principle irreducible to physics.

¹³ According to Quine in order to determine the ontological commitments of a scientific theory one has to regiment it in first-order predicate logic – which is an expression of Quine’s puritanism (on which see below), since in this way no property would ever be admitted in our ontology

¹⁴ deVries (2005; 2010). On the influx of Husserl on Sellars, see Moran (2012, 292-93).

world which, after separate scrutiny, [philosophers] must fuse into one vision” (Sellars 1962, p. 4). And, again like Husserl, Sellars (1963) also acknowledged that, genetically and methodologically, the scientific image of the world depends on the manifest image and that the normative concepts of the manifest image are not reducible to the descriptive ones that characterize the scientific image (see O’Shea 2007). It is from the ontological point of view, however, that the unilateralism of Sellars’s view becomes antithetical to the unilateralism of Husserl’s view. In fact, while Husserl was a realist about common sense and an antirealist about science, Sellars took the opposite stance.

According to Sellars, in the modern age the scientific image has justifiably acquired the monopoly on ontology, while the world of common sense has been shown to be unreal (and it is to be remembered that, on this, Sellars was stricter than Quine, since Sellars denied also the existence of abstract entities).

Sellars expresses this point through a now-famous neo-Protagorean motto:

Speaking as a philosopher, I am quite prepared to say that the common sense world of physical objects in Space and Time is unreal – that is, that there are no such things. Or, to put it less paradoxically, that in the dimension of describing and explaining the world, science is the measure of all things, of what is that it is, and of what is not that it is not. (Sellars 1956, 83; see also DeVries 2012).

Strict naturalism has inherited from Quine and Sellars a “puritanical” ontological attitude, as Stephen Stich (1996, 199) called it, according to which only scientifically acceptable phenomena are real. And this attitude explains the present multitude of naturalization projects that concern the features of the common sense world, such as persons, minds, tables, colors, qualia, free will, intentionality, normativity, and responsibility. Sadly enough, however, as Putnam (2004, 62) sardonically put it, “none of these ontological reduction gets believed by anyone except the proponent of the account and one or two of his friends and/or students.”¹⁵

In fact, many strict naturalists acknowledge that the alleged reductions of the features of the common sense world do not work. These philosophers have two possible ways out. Some (such as Paul and Patricia Churchland or the advocates of the so-called “Canberra plan”)¹⁶ take the radical step of dismissing the entire common sense view of the world as radically wrong. Others (the so-called ‘mysterians’) claim that we cannot give up beliefs as indispensable as those of freedom, responsibility, morality or consciousness; consequently, they conclude that the impossibility of naturalizing these notions is, and will always will be, an utter mystery for beings like us (McGinn 1999).

Both these proposals are generated by a certain background ideology that they share with all versions of strict naturalism – and New Realism consists, I

¹⁵ An interesting case in this regard is offered by Searle, who tries to reconcile the ontological puritanism typical of strict naturalism with a defense of an ontology of institutional facts, which in his opinion depend on collective intentionality. For a critical analysis of this view, see Smith (2003).

¹⁶ P.M. Churchland (1996); P.S. Churchland (2002); Nolan (2010).

would suggest, in the view that it is time to abandon this background ideology. As Tyler Burge wrote some time ago:

The flood of projects ... that attempt to fit mental causation or mental ontology into a ‘naturalistic picture of the world’ strike me as having more in common with political or religious ideology than with a philosophy that maintains perspective on the difference between what is known and what is speculated. (Burge 1993, p. 117)

Where common sense realism neglects science in order to solve the apparent clash of the two views of the world, strict naturalism is based on a dismissive attitude toward common sense that produces very dubious philosophical consequences. It is time to ask if there is a way of embracing a serious dual realism – one that would take seriously both the scientific and the common sense view of the world.

4. LIBERAL NATURALISM AS THE MOST PROMISING FORM OF REALISM.

In recent years some philosophers have developed new ideas in order to overcome the rigid, and unpalatable, alternative between common sense realism and scientific realism. In this light, very interesting proposals have come from forms of naturalism that are explicitly based on an egalitarian attitude in regard to science and common sense. It is in that direction that, in my opinion, we should look for the most promising new forms of realism.

In general, in the last years two main families of liberal naturalism have been developed. The first can be called *realist liberal naturalism* and its advocates attempt at revitalizing ontological pluralism without falling back into supernaturalism,¹⁷ the conception that claims the existence of entities that are incompatible with the scientific worldview.¹⁸ The second form can be called *quietist liberal naturalism* and its proponents – inspired by Wittgenstein’s quietism or by a naturalized Kantianism or for more substantial reasons – assume a quietist attitude in metaphysics and, consequently, tend to conceive of both the common sense and the scientific views of the world as indispensable, mutually irreducible, and categorically separated ways of conceptualizing the world.¹⁹ These approaches share the idea that strict naturalism illegitimately limits the scope of the term “nature” to the subject studied by the natural sciences. In this respect, one can extend Jennifer Hornsby’s view to all forms of liberal naturalism:

¹⁷ See Audi (2000) for a defense of the thesis that some forms of theism are compatible with the scientific worldview and therefore with naturalism.

¹⁸ See, for example, the proposals of Stroud (1996; 2006), Hornsby (1997), Smith and Ceusters (2010), Putnam (2012, section I; forthcoming), Baker (2013) defends an interesting emergentist conception, “Near-naturalism”, that is compatible with both liberal naturalism and supernaturalism. By adding an antisupernaturalist proviso to that conception, it could be appealing to many liberal naturalists.

¹⁹ See McDowell (1994; 1995), White (2007), and Macarthur (2008).

The world in which the mind is accommodated by [liberal naturalism] is naively natural; it contains the objects that we see and we act on; no peculiarly scientific method is required to have knowledge of it. (Hornsby 1997, 12)

However, since quietist liberal naturalism is by definition not concerned with the metaphysical issue of realism, this view is not of interest here.²⁰ On the contrary, the realist version of liberal naturalism, with its attempt to reconcile common sense and scientific realism in a non-Cartesian pluralist ontological perspective, is the most promising form of New Realism.

The tenets of realist liberal naturalism are:

- (i) *A liberalized ontological tenet*, according to which some real and non-supernatural entities exist that are irreducible to the entities that are part of the coverage domain of a natural science-based ontology;
- (ii) *A liberalized epistemological tenet*, according to which some legitimate forms of understanding (say, *a priori* reasoning or introspection) are neither reducible to scientific understanding nor incompatible with it;
- (iii) *A liberalized semantic tenet*, according to which there are linguistic terms that refer to real non-supernatural entities that do not form part of the coverage domain of natural science and are not reducible to those entities which do;
- (iv) *A liberalized metaphilosophical tenet*, according to which there are issues in dealing with which philosophy is not continuous with science as to its content, method and purpose.

Unsurprisingly, realist liberal naturalism is not left unchallenged.²¹ The first criticism that this view has to counter is that it is conceptually impossible, because there is no logical space between antinaturalism and strict naturalism. Ram Neta (2007), for example, wrote:

What if digestion, or respiration, or reasoning are natural kinds, their nature consisting simply in the mechanisms that enable them to occur? Is the liberal naturalist committed to denying this possibility? If so, then I confess I can see no good reason to accept Liberal Naturalism. And if not, then I confess I do not understand just what Liberal Naturalism is.

Put in this way, this criticism has the form of a dilemma: either liberal naturalism is not liberal enough (because it differs only superficially from strict naturalism) or it is not a genuine form of naturalism at all (that is, it is a form of non-naturalism in disguise). However, both sides of this dilemma seem ungrounded.²²

²⁰ For an important criticism of metaphysical quietism in general, see Fine (2001).

²¹ In the following, for brevity I will refer to realist liberal naturalism as “liberal naturalism” *simpliciter*.

²² Ram Neta’s dilemma is discussed more in depth in De Caro and Voltolini (2010).

Of course, if (like Goetz and Taliaferro 2008, 95) one presupposes that in order to be substantially different from strict naturalism a philosophical view should admit that the supernatural plays a causal role in the world, then it logically follows that liberal naturalism is not substantially different from strict naturalism. However, this way of putting the issue is clearly biased in favor of supernaturalism. If one takes a more neutral stance, then to see the deep differences between liberal and strict naturalism it is enough to look at the very different results that – because of their different presuppositions, methods and goals – the two versions of naturalism arrive at in dealing with concrete issues, from free will to morality, from consciousness to action theory, from the mind-body problem to the ontological status of secondary properties. In brief, while strict naturalists are committed to offering either reductionist or eliminationist analyses in regard to each of those issues, liberal naturalists will tend to take antireductionist stances (even though, it should be noted, they do not have to be antireductionist with regard to all those issues at the same time).

At this point, a strict naturalist could respond that, by expanding the scope of nature beyond the reach of the natural sciences, unavoidably the liberal naturalist paves the way to supernatural entities and explanations. In this perspective, beyond the reach of the natural sciences there is only the supernatural. Against this argument, the liberal naturalist could repeat that they refuse any appeal to the supernatural, both in ontology and in epistemology, but this defense is often contested. As a matter of fact, while the strict naturalist can easily delimit the space of the natural (and, consequently, that of the supernatural) by appealing to the scope of the natural sciences, the liberal naturalist seems to have a problem in this regard.²³

Let us take the case of mental causation. *Prima facie*, it may seem that if one claims that it is a real phenomenon but not explainable by the natural sciences, one is taking a supernaturalist stance. This, however, is not necessarily the case. Certainly, if one defends a strong form of dualism, according to which an immaterial mind intervenes in the natural world and breaks its laws, then one is *prima facie* a supernaturalist. But one can instead argue that the mind has causal powers that are neither reducible to nor incompatible with the laws investigated by the sciences of nature (Lowe 2006; Baker 2013, *passim*).

In general, liberal naturalists relax Quine's ontological criterion, since they do not think that one should only accept the existence of the entities that pull their weight in the best scientific theories of the natural world. In their opinion, one can also accept the entities that are implicit in out other sound and successful epistemic practices (such as common sense and the social sciences), as long as those entities are not incompatible with the natural science worldview.

Liberal naturalists conceive of nature as encompassing both the entities accepted by the natural sciences and those accepted by common sense. In this respect, one can quote Stroud (2006, 350), when he wrote: “[We] need a conception of nature that includes a whole world of enduring bodies in space and human bodies and human actions in interaction with them and with one another” –

²³ I thank Matteo Morganti for bringing this problem to my attention.

and add to this also other non-supernatural phenomena that cannot be explained by the natural sciences, such as phenomenological properties and debts.

At this point, a strict naturalist could reformulate his charge and say that, by refusing the possibility of a totalizing mechanistic explanation of all reality, liberal naturalism, if different from supernaturalism, is still too inclusive, since it makes room for some obviously non-naturalistic views, such as cultural relativism and deconstructionism. And this would be something that not even the advocates of those views would be happy with.

However, in taking a doubly realist stance that embraces both the scientific and the common sense view of the world, liberal naturalism is immune from this charge. This view in fact claims that reality limits the scope of the legitimate interpretations of reality by determining the objective conditions of truth of our scientific and common sense judgments. Certainly, our judgments, both in science and common sense, are fallible, and always will be; but that does not mean that there are no objective standards of truth, as claimed by cultural relativism, deconstructionism and the like (Boghossian 2006).²⁴

Another criticism against liberal naturalism grants its conceptual legitimacy but claims its implausibility. This criticism has been expressed in different ways. One is based on the ‘burden of proof argument’, claiming that strict naturalism is the default naturalistic view. In this perspective, liberal naturalists should prove that the natural sciences cannot accommodate some parts of the world, as we know it – particularly of the human world – otherwise theirs would not be a legitimate metaphilosophical option. That is, the liberal naturalist has to prove that some of the features of the world are both ineliminable and in principle irreducible to features acceptable to natural science (see for example, Macdonald 2006, 231).

This argument is also unconvincing however. First of all, it is unfair that the strict naturalist, who by definition refuses the possibility of *a priori* arguments, asks the liberal naturalist for an *a priori* demonstration that some features of the world are both ineliminable and irreducible to its scientific features. Moreover, by the same token, the liberal naturalist could argue that the burden of proof is on those who claim that the recalcitrant features of the world (free will, consciousness, debts, and so on) can be reduced to the entities accepted by the natural sciences or explained away.

Another way of putting the charge of implausibility against liberal naturalism is based on the so-called ‘Great success of science argument’.²⁵ According to this argument, starting with the scientific revolution, the natural sciences have progressively explained an astonishing quantity of previously very mysterious-seeming phenomena, making it possible for us to predict and control many of them at will. Therefore, it is rational to infer that also the problems of agency, consciousness, morality, and so on one day will also be solved (or, at least, could in principle be solved) by the natural sciences.

²⁴ In regard to ethics, for example, Putnam (2004) shows how it can be objective, irreducible to non-normative forms of knowledge, and non-supernatural (he takes a stance that is epistemically and semantically realist and ontologically antirealist).

²⁵ This argument is discussed and criticized in more detail in De Caro and Macarthur (2004a).

Not even this argument sounds persuasive. First of all, it is far from obvious that the induction on which this argument is based is adequate. Moreover, as has been noted (for example in Crane and Mellor 1995), it is unclear to what scientific theories the Great Success of Science argument refers. Certainly, they cannot be present-day theories, since these are unable to solve the problems of agency, consciousness etc. But who knows what forms the theories that in principle *are* able to explain those problems might have (granting that such theories do indeed exist)? Should we then not confine this kind of speculative reasoning to the (already crowded) realm of overambitious philosophical phantasies? Moreover, this argument begs the question when it is used against liberal naturalism. The latter view is based on the idea that it is rational to believe that some important features of the world are ineliminable and not reducible to the features accounted for by natural science: just saying that they will be eliminated or reduced because science is intrinsically able to do that sounds very much like a mere repetition of the thesis that liberal naturalists contest.

Finally, another way of denying plausibility to liberal naturalism consists in insisting on the indubitability of ontological and epistemological monism (compare P.M. Churchland 1996; Schaffer 2014). From this point of view, all the features of the world that presently look ineliminable and irreducible to features acceptable to natural science are in principle either eliminable or reducible. At the scientific level, this insistence on monism may sometimes be justified as a methodological ideal. However, if one takes monism as a restriction on how reality should be, then not only is it at odds with common sense, but also with scientific practice, since today pluralism is a very common view among the natural sciences,²⁶ and it goes without saying that the monistic cause becomes even less promising if one takes into account the social sciences as well. Indeed, the idea that pluralism represents a menace for the scientific view of the world is a heritage of an obsolete positivistic view that it would be time to forget.

The attempt to reconcile realism about common sense and realism about science in a naturalistic spirit thus survives all the charges that have been raised against it. The philosophical marsupials may have finally found a new luxurious continent to explore.²⁷

²⁶ See Kellert *et al.* 2006. According to an authoritative recent survey, even the possibility of reducing chemistry to physics (traditionally the best showpiece of reductionism) is today widely contested: “Most philosophers of chemistry think that a stronger conception of unity [between chemistry and physics] is mistaken. Most believe that chemistry has not been reduced to physics nor is it likely to be” (Weisenberg *et al.* 2011).

²⁷ My thanks to Robert Audi, Lynne Baker, David Macarthur, Maurizio Ferraris, Andrea Lavazza, Massimo Marraffa, Matteo Morganti, Hilary Putnam, Barry Smith, Karsten Stueber, Alberto Voltolini, Stephen White for many useful discussions on these issues. I also thank the audiences of the talks I gave the University of Notre Dame, Oxford University, University of Massachusetts at Amherst, College of the Holy Cross, University of New Hampshire, Milano San Raffaele, Paris IV-Sorbonne, and Torino for the very useful comments on previous versions of this paper.

REVISITING THE MIND-BRAIN PICKLE: SWINBURNE'S EVENT DUALISM OF SUBSTANCES AND PROPERTIES

NYTHAMAR DE OLIVEIRA

1. In this brief paper, I would like to confine myself to offering a few comments on Professor Richard Swinburne's seminal contributions to the philosophy of mind and language, and to the philosophy of neuroscience, especially in his work *Mind, Brain, and Free Will*, and more particularly on his recasting of the mind-brain problem, his critique of physicalism, and his defense of a variant of substance/property dualism. Swinburne's major guiding thesis comes down to asserting that physical, material events and conscious, mind events (such as beliefs, desires, thoughts, and sensations) are not identical, so that "the mental world" cannot be deemed "fully deterministic." (Swinburne, 2013, p. 202)

Since I believe that many neuroscientists and philosophers of mind nowadays who embrace physicalism are motivated by their correlated refusal of substance and property dualisms (I am thinking here of António Damásio and Jesse Prinz) without succumbing to determinism or to a reductionist account of naturalism (say, like the Churchlands' eliminativist program), I would like to revisit here some of the features of this classical problem which resists any clear-cut solution, in light of Swinburne's recasting of substance and property dualisms. (Churchland, 1984, 1995, 1996) I am thus assuming that the brain-mind problem is analogous to the nature-nurture pickle, in the sense of a predicament like a chicken-egg question, not so much to sort out which one comes first but which conditioning is more basic and takes primacy over the other. Even though it could seem trivial to realize that there are no mental events without brain processes, just like neurobiological evolution (nature) would intuitively seem to be prior to social, cultural evolution (nurture), the fact that human cultural and civilizational processes have transformed nature throughout the centuries and millennia just attest to the difficulties involved in the brain-mind pickle. (Prinz, 2002) In a US idiomatic expression, if you are in a pickle, you are in a difficult position, or have a problem to which no easy answer can be found –so that happens to be case with the brain-mind problem, following a neuroscientific turn in the recasting of the body-soul problem.

¹ Associate Professor of Philosophy, PUCRS; Brazilian National Research Council (CNPq) Fellow. I would like to express my gratitude to Professor Richard Swinburne for his highly stimulating lectures and for giving us the honor of his visit to our graduate program in Porto Alegre.

Professor Swinburne starts from the basic assumption that mental events (consisting in the instantiation of mental properties – sensations, thoughts, purposes, desires, and beliefs) are distinct from physical events (such as brain events), although in causal interaction with them. As expected, ontology is the starting point for his mapping of concepts and theoretical framework: “the whole history of the world can be told with our familiar system of categories: substances, properties, and times. I understand by a substance a particular concrete object: my desk, that person, the photon (particle of light) emitted from this light source which landed on this screen, and so on. Substances may have other substances as parts. My desk has its drawers as parts of it; and it can exist (it is logically possible) independently of all other things of its kind (i.e. all other substances) apart from its parts; and those parts have very many electrons, protons, neutrons, etc. as their parts. Substances exist all-at-once. Whenever they exist, they exist totally.” (Swinburne, 2013, p. 4) When I first read these lines, I couldn’t help drawing comparisons to Edmund Husserl’s parts-whole formal ontology, namely, recasting Leibniz’s monadology as an alternative third way, so as to avoid both Descartes’s substance dualism and Spinoza’s holistic monism. I think, *en passant*, that one might spot a neurophenomenological deficit in both normative and naturalist theories that fail to account, respectively, for neural correlates in first-personish reconstructions of social action (as in most critical-theoretical approaches of second and third generations of the so-called Frankfurt School) or in naturalist theories of sociality that miss the irreducibility of semantics to syntax and the former’s pertaining to phenomenal consciousness of aboutness and what’s-it-likeness (as in the Churchlands’s eliminative materialist research program). Although I won’t pursue this point here, we may think of Habermas’s critical, normative reconstruction of subjectivity as he conceives of cognitive- and moral-psychological development as a rationalization of the structures of consciousness, as the conventional perspective can be replaced by the postconventional perspective, for instance, when adolescents consciously grow and reflect upon their own justifications of moral, normative principles.(De Oliveira, 2014) In the final analysis, it is quite understandable that neuroscientists overall and neuropsychologists would refer to physical or natural ontology as they deal with real, natural beings, their properties and events that can be described and explained as physical phenomena –without any resort to supernatural or metaphysical discourse. For physicalists, materialists, and naturalists it suffices to take ontology as the real, phenomenal realm of beings, entities, phenomena, and events as they appear, come into being or exist, necessarily, possibly or contingently, very much as traditionally and broadly conceived, as the study of what there is in the sense of real, objective existence. Naturalists like John Searle have shown, however, that epistemic and ontological takes on objectivity and subjectivity are not as straightforward for philosophers, say, when contrasting events or phenomena studied by so-called “hard sciences” with the social reality or institutions that are examined by social scientists, precisely because of the irreducibility of the phenomenal, self-conscious perspective of the first person who experiences pain, feels cold or exchanges goods with other human beings in their social

dealings.(Searle, 1984, 1995) In this sense, it can be argued that the social, intersubjective dimension of human selfhood, in its correlated capabilities of conscious reflexivity and autonomy, are the best way to account for the limits that humans impose on the treatment-enhancement divide (say, in neuroethics and bioethics), as the pursuit of a good life and human flourishing addresses questions of how we should lead fulfilling lives, not only to ourselves (self-fulfilling lives) but also to others (sociability). It has been my working hypothesis that by calling into question whether social, cultural conditioning can be actually undermined by neurobiological conditioning alone, it remains to be seen how positive social feedback drives people to interact on social media and, the other way around, whether use of social media ends up changing the way positive social feedback is actually processed by the brain. Since most social thinkers tend to identify sociality with intersubjectivity (for instance, the fact that shared beliefs or social norms are common to individuals belonging to the same social group or set of individuals), one may think of social institutions (broadly conceived so as to include not only the state, governmental, political, economic, and legal structures, but also the family, civil society, organizations, associations, and social grouping of all sorts) very much by analogy with the way one learns how to function in a natural language (sharing a grammar, phonetics etc), usually without paying much attention to it (esp. mother tongue for native speakers, as analogous to the way one has been socialized into being Amish in Amish country)

2. At any rate, the point being made here is that, according to Swinburne, mental events consist in the instantiations of properties in the immaterial substance, which has been traditionally termed soul, *anima*, as the seat par excellence of reflexivity, interiority, and self-consciousness. Accordingly, humans and nonhuman animals alike (esp. “higher” animals) are said to consist of two parts, the essential part (the soul) and a contingent part (their body). Since Aristotle, we think of humans as peculiarly distinct vis à vis the higher animals because of the former’s ability to speak and reason (*logon echon*) logically and morally, hence the very conception of free will has been an integrated system of beliefs and desires. But we must unpack how these categories of substances, properties, and times concur to pick out mental events and brain processes. According to Swinburne, “an event as either some substance (or substances, or event or events) having a certain property (more formally, the instantiation of a property in some substance or substances, or event or events) at a certain time, or the coming into existence or the ceasing to exist of some substance at some time.” (Swinburne, 2013, p. 6) Once we agree that substances, properties, and events are basic constituents of the world and once we give in to his principle of credulity — i.e., “what seems to us to be the case probably is the case, absent any counter-evidence” (Swinburne, 2013, p. 42f) —, Swinburne proceeds to persuade us that the soul *qua* substance is, after all, the essential part of our being insofar as it can exercise causal power over the body, and more importantly, free from deterministic causes in such a way as to render us morally responsible for what we do. It seems quite problematic to assume that by

confining ontology to existence (as opposed to, say, being and modes of being), Swinburne would succeed in avoiding Platonic realism and its correlated semantics, according to which all nouns do refer to existent entities (in a transcendent realm of forms), or other variants of general metaphysical or ontology that might undermine his contention that substances have properties, as opposed to one single substance (monism). Hence, when Swinburne arguably points out that “philosophers and scientists have made claims about what is ‘possible’ in this area, such as ‘it is not possible for a person to exist without a body’ or ‘necessarily all mental events supervene on physical events’” to add that “whether that is true depends on what is meant by ‘possible’ and ‘necessarily,’” he seems to be committed to a particular semantic-ontological framework. (Swinburne, 2013, p. 4ff) Thus, in order to account for the soul’s interaction on the body and material things, Swinburne strategically adopts the following extended notion of supervenience: “A-substances supervene on B-substances iff necessarily for every A-substance x there is a B-substance y, such that necessarily if y exists x exists.” (Swinburne, 2013, p. 21) Precisely because of his ontological commitments, Swinburne strongly argues against any “restriction of the mental to the sensory” (Swinburne, 2013, p. 98), citing David Chalmers’s property dualism (akin to his own version of event dualism), according to which there are “both physical and non-physical features of the world. The falsity of logical supervenience implies that experience is fundamentally different in kind from any physical feature.” (Chalmers, 1996, p. 124) One of the reasons why Swinburne’s case for dualism seem to fail to deliver the normative grounds promised by his otherwise highly original and critical account of property dualism is, perhaps, to be found in its lack of commitment to what Chalmers dubbed “the hard problem of consciousness.” (Chalmers, 1995) As R.D. Ellis put so well, “the main point of the hard problem is that, even if we could discover the ‘neural correlates of consciousness,’ we still would not have answered the ‘harder’ question: Why do those physical events exhibit the property of consciousness, whereas other physical events do not?” (Giordano and Gordijn, 2010, p. 66) Even though these remarks seem to address the naturalist horn of the dilemma, the second horn turns out to render problematic the very meaning of normativity caught between the subjectivism of first-personal accounts and the absolutism of third-person accounts (esp. absolute principles and divine command theories as in the Euthyphro dilemma).

3. To be granted, Swinburne carefully distinguish beliefs and intentions as continuing mental states that do not by themselves entail any physical events involving the believer or agent, as opposed to, say, desires and dispositions to do actions: “In that they exist over periods of time during which they are totally absent from my consciousness, they are clearly continuing mental states and not—like intentions in action—conscious events.” (Swinburne, 2013, p. 83f) Now, physicalists like Damásio and Prinz have convincingly argued that reason, emotions, and decision-making processes can be articulated in terms of empirical and philosophical language, in that cognitive feelings and a reflective level are integrated with noncognitive features of emotions and preferences,

particularly the so-called “primary emotions” and “gut reactions.” For one, Damásio has decisively contributed to ongoing interdisciplinary research in cognitive sciences, neurophilosophy, neurobiology of mind and behavior, particularly at the crossroads of emotions, decision-making, memory, communication, creativity, and consciousness as neurophysiological phenomena that call into question reductionist approaches. Indeed, the publication of his *Descartes’ Error*, in 1994, started off a real turning point not only in neurology, psychiatry, neuroscience, and cognitive psychology, but also in the philosophy of mind and language, as it undertook a radical critique of Cartesian dualism, opposing dichotomies of soul and body, brain and mind, reason and emotion. (Damásio, 1994) Since the 1950s and 60s, research in neuroscience has already shaken apparently insurmountable problems in various models of dualism and of several others that have emerged in the following decades, with alternative proposals to patterns of behavior conditioning (behaviorism), theories of identity (between mind and brain), the physical states of the brain (physicalism) and their causal roles and functions in a complex economy of internal states, mediating sensory data inputs and behavioral outputs (functionalism), as well as the materialistic reductionisms that supposedly eliminate folk psychology and normative accounts that allude to psychological states (eliminative materialism). Damásio’s work fostered thus a fruitful dialogue between neuroscientists and philosophers of mind, especially within neurophilosophy and cognitive sciences, as attest seminal works by Searle, Gazzaniga, and Prinz. Of particular concern is their recasting of the “social brain” problem, as Damásio and Prinz assume that the philosophical underpinnings of cognitive and moral decisions are at the center of discussions about human nature, in that self-conscious morality-*cum*-sociality evolves as one of the elements that distinguish humans from superior primates and other nonhuman animals. Moral decisions occupy, after all, a central place in defining the human being, at the heart of decisions that define us in relation to cultural, social problem-solving, relationship issues, and personal and political choices that ultimately help us set the “self” in everyday relations to ourselves and to the others and within a particular milieu. Damásio establishes thus the correlation between practical reason and emotion, combining the awareness notion of decision-making and planning at different time scales, creating possibilities of interaction with the environment and the selection of courses of action, with all processes and steps interconnected. Damásio manages thus to articulate the social, intersubjective, and neurobiological processes that explain the evolution of the human brain and the emergence of consciousness, the “I” –as a first-personish self, a reflected-upon “me” and third-personish accounts of other selves—, memory, language, subjectivity, and their representations and creative constructions and carriers of meaning. According to Damásio, “Both basic homeostasis (which is nonconsciously guided) and sociocultural homeostasis (which is created and guided by reflective conscious minds) operate as curators of biological value. Basic and sociocultural varieties of homeostasis are separated by millions of years of evolution, and yet they promote the same goal—the survival of living organisms—albeit in different ecological niches. That goal is broadened, in the case of sociocultural

homeostasis, to encompass the deliberate seeking of well-being. It goes without saying that the way in which human brains manage life requires both varieties of homeostasis in continuous interaction. But while the basic variety of homeostasis is an established inheritance, provided by everyone's genome, the sociocultural variety is a somewhat fragile work in progress, responsible for much of human drama, folly, and hope. The interaction between these two kinds of homeostasis is not confined to each individual. There is growing evidence that, over multiple generations, cultural developments lead to changes in the genome." (Damásio, 2010, p. 31)

4. Indeed, since Gazzaniga (1985) formulated the problem for the first time in the 1980s, the "neural substrates" of social behavior and cognition have not yet been completely understood. Moreover, studies in humans and other primates have revealed different neural structures that play a decisive role in the construction of social behavior: the amygdala, the ventromedial frontal cortices and the right somatosensory cortex, among other structures, which seem to mediate perceptual representations of socially relevant stimuli. These studies made it possible to develop the Social Brain Hypothesis, according to which the restrictions on the size of the social group arise from the ability of information processing in the brain, especially among primates, so that the neocortex eventually play an important role in social evolution that leads us to present complex sociality. Thus the Dunbar number was first proposed in the 1990s by British anthropologist Robin Dunbar (1998), who found a correlation between primate brain size and average social group size.

However, even such a proposal raises a number of interpretations on how this relationship is mediated. For Dewey, who influenced the normative-reconstructive approach of social thinkers like Rawls, Habermas, and Honneth, thought is necessarily symbolic and all social symbolism is necessarily, therefore, the mind is always already social: there are sources of expertise outside the individual, insofar as we have to live from birth to death in a social world of people and artifacts, which is largely the result of what has been done and transmitted from previous human activities in concert --in cultural traditions, through linguistically and socially mediated contexts of meaning. When this fact is ignored, experience is treated as if it were something that happens exclusively within the body and mind of a lonely individual or disembodied self. According to Dewey, experience does not occur in a vacuum, but it always presupposes an intersubjective externality to an individual, which gives rise to the very experience itself within a social world. Certainly, not all sociality can be reduced to brains, not their conceptualizations can be socially determined. According to social epistemology, the emphasis on the primacy of emotions and the importance of common notions are not always equally crucial to characterize the formation of knowledge, agreement and disagreement between epistemic peers in decision-making groups. The social dimension that is being emphasized in discussions of social intellect, culminating with the notion of Machiavellian intelligence and its presence in the world of primates, is the individual's ability to interact successfully with the social groups in order to predict and manipulate

the behavior, making and breaking promises, and so on. The energy requirements of such a complex situation are ultimately presented as responsible for the large size of the primate brain, so that some evolutionary anthropologists and researchers in related fields postulated the hypothesis of Machiavellian intelligence and the social brain hypothesis. Moreover, the concept of social brain is not reducible to the individual manifestations of a social world around us because the “brain architecture” reflects rather than forms its social organization, language and culture. It is against such a complex semantic context that can be investigated the processes of moral decision-making and ethical implications that materialize in everyday life and social media, as measured in neuroimaging experiments. Beyond the culturalist, rationalist and modular approaches to language, this research will thus help us figure out how language and cross-cultural identities (including gender, ethnic, social, political etc) function in social interactions comprising diverse fields such as pragmatics, neurolinguistics, and neurosemantics.

So the first set of questions that addresses Swinburne’s recasting of substance/property dualism can be thus formulated: since a phenomenological notion of an embodied mind or of a minded body does not entail an endorsement of some kind of Cartesian materialism, as if we were simply getting rid of the soul (or the mind, for that matter), how about embracing a more explicit phenomenology of mind? As Professor Swinburne interestingly remarks in his own criticisms of misleading, reductionist interpretations of Libet’s experiments: “In other cases it does not seem to us that we are choosing without being caused to choose as we do, and so we should not believe that we are then making an uncaused choice. The phenomenology of deciding between rival possible actions, ones which are not determined by our mental states (our existing desires and beliefs with their relative strengths), is so different from the phenomenology of doing the everyday things we do intentionally, that we should expect the underlying brain processes to be similarly different.” (Swinburne, 2015, p. 201f.) After all, an agent has free will, as Swinburne goes on to assert, “insofar as the agent acts intentionally without their intentions being fully determined by prior causes.” Or as Swinburne put it, “having an intention in making” such and such is equivalent to ‘having an intention which the subject believes causes him or her to make them’ rather than—as ordinarily—‘having an intention which causes the subject to make them’.” (Swinburne, 2015, p. 80f.) The irreducibility of first-personish accounts of beliefs and actions in response to phenomenological, normative challenges (esp. when dealing with intentionality, aboutness, and consciousness) that avoid trivial conceptions of normativity and naturalism might help us avoid the deterministic attempts to reduce the sense of normativity, say, as in Jennifer Hornsby’s conception of Naive Naturalism, according to which in order to avoid both physicalist and Cartesian claims about the mind-body problem, we ought to return to common sense and folk psychology as they implicitly endorse normative and first-personish beliefs. In a nutshell, is it the case that supervenience of moral properties on non-moral properties must be analogous to supervenience of substances, as Swinburne extended Kim’s conception so as to account for natural, physical phenomena

without resorting to deterministic approaches such as the identity theory? That being the case, how does property or event dualism avoid the criticisms raised against functionalism? As Swinburne saw the problem, functionalists claim (to use his own terminology) “that what makes any property a property of a kind which [Swinburne has] called ‘pure mental property’ is that events with that property have a certain function in a person’s life of thought and behavior, and in particular tend to have certain kinds of causes and effects (in or outside the brain).” (Swinburne, 2015, p. 94) Granted, I can see that this strategy will be very helpful in keeping moral normativity separate from naturalism or physicalism. As Swinburne put it so bluntly, “Moral beliefs as such, I suggest, like all value beliefs and unlike other beliefs, motivate us. I could not believe that some action was really morally good to do (as opposed to being what other people call ‘morally good’) and yet not see myself as having a reason for doing it.” (Swinburne, 2015, p. 178)

5. Now, still relating to property supervenience and causation, we might raise the question of how to interpret the correlations between mental and neural phenomena discovered by brain science and psychology since their beginnings. Brain research suggests that there is a high degree of covariation between mental states and brain states. This view, however familiar, raises more questions than it tends to be aware of, and we may as well pick out just three of them, following Prinz (2002, p. 71ff; 2004a; 2012, p. 168 f):

- (1) If causal relations are at all possible in a transphysical context how do they have to be conceived?
- (2) How far are “passive” mental events causally dependent on brain processes?
- (3) How far are “active” mental events causally relevant to brain events?

Whereas the first question is largely a challenge to philosophical analysis, the two other questions are a challenge, and an opportunity, for a coordinated effort of all contributing disciplines, including the neurosciences. Since Swinburne admits that instead of an event dualism, it could be that the public world (not merely our description of it) contains some other dualism (an ‘aspect’-dualism, for example) which turns out to be just a different way of describing the same feature of the world as does ‘event-dualism’. Now, could we stretch this to go so far as to say that perhaps event dualism allows for a perspectivism that avoids an ontological dualism, like in Kant’s noumenal-phenomenal dualism understood as transcendental perspectives rather than ontological realms? Isn’t the case that Kant himself also made room for agent causation, as opposed to inanimate events (e.g. the motion of billiard balls), allowing for the first-person account of autonomous, self-legislating selves that cannot be reducible to third-person, descriptive accounts precisely because of their peculiar practical faculty to initiate a series of events in nature? Indeed, for many physicalists, the physical is sufficient to generate the mental and that a further causal contribution (say, of something supernatural) is not called for. Although the causal sufficiency of the

physical cannot rule out a supernatural influence categorically, such an influence would be redundant. It would not be needed to explain the existence and functioning of the mind. All in all, one cannot speak of naturalist normativity or normative naturalism without a certain embarrassment. And yet, as over against traditional conceptions that regard naturalism as merely descriptive, as opposed to prescriptive accounts of normativity, it has become more and more common nowadays to challenge such a clear-cut division of labor, as naturalists like Millikan (1989) assign normative force to the biological concept of function and normativists like Korsgaard tend to assume that human psychology is naturally normative: “whatever confers a normative status on our actions – whatever makes them right or wrong – must also be what motivates us to do or avoid them accordingly, without any intervening mechanism.” (Korsgaard, 2010, 16) To be sure, both views could be regarded as simply recasting the externalist-internalist debate over the problems of teleology, intentionality, motivation and carrying out an action supposed to be moral. Once again, Damásio’s integrated views of emotions and feelings not as “intruders in the bastion of reason” but enmeshed in its networks, for worse and for better, are revealing: “The strategies of human reason probably did not develop, in either evolution or any single individual, without the guiding force of the mechanisms of biological regulation, of which emotion and feeling are notable expressions.” Accordingly, empathy is a highly flexible, context-dependent response to these networks, ultimately leading to cooperation and the evolution of social norms, especially fairness norms. Damásio evokes thus the process of a sociocultural homeostasis so as to refer to the social and cultural imbalances allowing for the detection of an imbalance at a high level of a conscious brain-mind in the stratosphere and not in subcortical level. Damásio’s takes on emotions and feelings within an integrated 4EA-view of cognition (embodied, embedded, extended, enactive, and affective), very much like Prinz’s, allow for a homeostatic understanding of the development of moral rules, laws, and justice systems (very much like an effect of a wide reflective equilibrium), as a response to the detection of imbalances caused by social behaviors that make endanger individuals and the group. The cultural devices created in response to the imbalance aim to restore the equilibria of individuals and the group. So people are capable of social cooperation and empathy, but they can be also callous, indifferent and socialized into *schadenfreude* (finding pleasure in others’ pain) –the social, cognitive, and neural mechanisms underlying empathy and that may help to alleviate humanity’s deepest tragedies and facilitate its greatest triumphs. So this intricate connection of the body to emotions is related to homeostasis, which can be rethought of as the machinery regulating life that also has to do with the development of culture. This development manifests the same goal as the form of homeostasis. It reacts to the detection of an imbalance in the process of life and seeks to correct it within the limits of human biology and the physical and social environment. The contribution of economic and political systems, as well as, for example, the development of medicine, are a response to functional problems that occur in the social space and require a correction in this space, so that will not undermine the regulation of vital individuals that constitute the

group. We come thus full circle within a broad understanding of wide reflective equilibrium, in sociocultural homeostatic and social-ontological terms, allowing for intersubjective and linguistic interactions and co-constitution of meanings.

6. Last but not least, since Professor Swinburne also pursued theology (1959-60), besides his undergraduate (1954-57) and graduate (1957-59) studies in philosophy at the University of Oxford, one wonders what he makes of the ongoing science wars and evolution wars in the US, especially those opposing the scientific community and fundamentalist and conservative Christian theologians and believers. As we all know, this creationism-evolution debate is not a real problem in Catholic and moderate, liberal protestant theology (or progressive, reform Judaism for that matter). Now, how, within Professor Swinburne's standpoint, the principles of credulity and testimony could be evoked to assess the belief, say, in the inerrancy of the Bible when dealing with creationism and miracles? Would they go so far as to follow Alvin Plantinga in holding that since belief in the theist, personal God is properly basic, then it would seem that belief in inerrancy would be, within the circumstances of Christian faith, a properly basic belief as well? I am raising these questions, out of curiosity but also with a view to testing the coherence of Professor Swinburne's dualism, as many issues relating to the composition of the Hebrew Bible and of the New Testament have led to conjectures and documentary hypotheses that rely on probability, for instance, that there is new evidence to assert, nowadays, that it is more likely (probable) that the Torah was composed much later than sooner (conservative traditional chronologies dated back to 1,200 BCE) and that much of the whole Hebrew Bible was written from the seventh through the 5th century BCE (according to many researchers, such as Israeli leading archaeologist Israel Finkelstein). So just like creationists who stick to a Young Earth hypothesis (between 6,000 and 10,000 years) to oppose the Big Bang (over 13.7 billions, with the Earth's age estimated in about 4.5 billion years), conservatives and fundamentalists still refuse to accept scientific contributions (e.g., archaeology and innovative methods for dating) in their own handling of Scriptures. Couldn't the principles of credulity and testimony, in this case, turn out to be quite misleading? After all, substance dualism could easily fall back into some subtle Manicheist doctrine of supernatural powers intervening in the natural cosmos, just like property dualists could still hold that some of our mental states have immaterial properties, even though we ourselves cannot be solely identified with immaterial souls wholly distinct from our bodies and natural properties.

In my own ongoing interdisciplinary research program in normativity and naturalism, I have been arguing for a mitigated conception of social constructionism that avoids both versions of dualism and still allows for a normative reconstruction of the so-called social brain hypothesis, so that phenomenal consciousness and first-personish accounts, including beliefs, desires, intentions, and propositional attitudes, cannot be ultimately eliminated by physicalism. By pursuing an interdisciplinary research in the philosophy of neuroscience, neuroethics, and social neurophilosophy, especially by focusing

on the relation between naturalism and normativity, one might avoid the reduction of either to the other, by stressing the inevitability of bringing in the two other poles of the semantic correlation whenever dealing with ontology, language, and intersubjectivity. As Jesse Prinz's takes on transformation naturalism and concept empiricism allow for an interesting rapprochement between social epistemology and critical theory, his critical view of both *naturism* (i.e., reducing the nature-nurture pickle to the former's standpoint) and *nurturism* (conversely reducing it to the latter) not only successfully avoids the extremes and reductionisms of (cognitivist) rationalism and (noncognitivist) culturalism –such as eliminative materialism and postmodernism—, but turns out to offer a better, more defensible account of social epistemic features and “social pathologies” than most analytical, social epistemologists (e.g. Goldman, Parfit) and critical theorists (Habermas, Honneth) have achieved thus far. I have argued that Damásio, Prinz, and Searle, among others, have succeeded in showing that the social brain rather than the solipsist mind is what must ultimately account for a scientifically informed theory of consciousness, as mental representations of a given stimulus located at an intermediate level of processing become conscious through attention. The semantic-ontological correlation comes thus full circle vis à vis its networking with language and subjectivity. As Prinz felicitously put it in his neoempiricist, reconstructive theory of emotions: “Moral psychology entails facts about moral ontology, and a sentimental psychology can entail a subjectivist ontology.” (Prinz, 2004b, p. 8) After all, the descriptive and experimental dimensions of most experiments fail to provide for such a moral justification, insofar as causality or causation cannot be taken for granted or satisfy ought-like normative claims –since not every correlation turns out to be causal.

7. Neuroethics, as I have argued elsewhere, deals with bioethical, moral problems both in abstract, theoretical terms (such as in metaethics and normative ethics, for instance, to define what is good and what selfhood is all about) and in practical, concrete terms (applied and experimental ethics), especially related and informed by the empirical sciences and recent findings in neuroscience. (De Oliveira, 2013, p. 84) Neuroethics deals precisely with this intersection of possible, imaginable uses of neurotechnologies and their moral acceptability, desirability, and permissibility: when is it permissible to alter a person's psychological conditions, dispositions, memories, to the point of influencing her personality traits or “reading” her mind? What can neuroscience tell us about free will, self-control, self-deception, conditioning mechanisms and the very justification of moral paths to be adopted by one individual or social groups? What neuroenhancement and neurotechnological interventions are morally acceptable and appropriate to be adopted in public health policies and legislation? To be sure, what is properly fearful and innovative in brain design enhancement is the ability to change something inherent in “human nature,” especially its genetic configuration and the implications of such changes. One might think of good examples in the neuropharmacological research and drug industry –both for enhancement and therapeutic purposes— as neurotransmitters,

such as serotonin, and hormones, such as oxytocin, have been manipulated with the aim of reducing anxiety or stimulating empathy and social engagement. That would be quite different from using neurotechnologies and nanotechnology to change or manipulate the human genome itself, while seeking some cognitive and life-quality enhancement, in case it entailed some form of liberal eugenics or social Darwinism. Mutations and genetic manipulation itself would not *per se* be morally questionable, but the way they could be managed and implemented, in that they could compromise fundamental bioethical principles such as individual freedom, social justice, non-maleficence, and beneficence arising from particular cases. What one learns from neuroethical theories and conjectures, in the last analysis, is that moral dilemmas are not solved once and for all by simply resorting to neurotechnologies or to neural correlations supposedly establishing causal relationships, as if all desirable effects could be obtained by the appropriate changes in the causes without damage or risks –to the patients themselves or to third parties. Indeed, one of the first contributions of neuroscience to social cognition is how the decision to take action in relation to a moral dilemma (say, in the classic dilemma of the runaway railway trolley) is associated with additional recruitment, as cortical, neural networks are associated with the need to exert cognitive control at a given moment, making impossible to reduce a decision to an automatic or procedural process. As studies in humans and other primates have revealed, different neural structures play a decisive role in the construction of social behavior and the so-called social brain: the amygdala, the ventromedial frontal cortices, and the right somatosensory cortex, among other structures, which seem to mediate perceptual representations of socially relevant stimuli, being reflexively sedimented in social, cultural codifications. In effect, the restrictions on the size of the social group arise from the ability of information processing in the brain, especially among primates, so that the neocortex eventually plays an important role in social evolution that leads us to our present complex sociality. The “Event-Feature-Emotion” complex or EFEC developed by Jorge Moll *et alii* suggests that moral, cognitive phenomena emerge from the interplay between three main components associated with the recruitment of specific brain centers: knowledge of structured events (contextual representations in prefrontal regions), social traits and functional features (stored in the temporal cortex, such as perceptual memories), and central, basic emotional states, such as aggressiveness, sexual arousal, attachment, and sadness (represented in limbic and paralimbic structures). Models such as EFEC can generate hypotheses about the neural bases associated with different moral dilemmas from what might be the motivation and cognitive processes that underlie the decisions made. These hypotheses can be tested from the association of specific situations and dilemmas with specific networks whose functions (say, providing the basis for central emotional states) are well established. The social implications of this neuropsychological complex have been explored in neuroethics, neurolaw, bioethics, and applied ethics overall: “Moral cognitive neuroscience researchers have developed innovative paradigms for the scientific exploration of unique forms of human social behaviour.”(Moll, 2005, p. 801f) Thus an

interdisciplinary research project on “Social Media and Decision Making: Reason and Emotion in Social Relations” has been implemented at the Brain Institute in Porto Alegre, Brazil, revisiting the so-called problem of the social brain from an intercultural, multidisciplinary perspective, including experiments in partnership with counterparts in Norway (University of Bergen) and the U.S. (University of Miami).² According to social epistemology, the emphasis on the primacy of emotions and the importance of common notions are not always equally crucial to characterize the formation of knowledge, agreement and disagreement between epistemic peers, and decision-making in social groups. In effect, the social dimension that is often emphasized in discussions about the social intellect, culminating with the notion of a Machiavellian intelligence and its presence in the world of primates, is the individual’s ability to interact successfully with social groups in order to predict and manipulate the behavior, the making and breaking of promises, and so on. The energy requirements of such a complex situation are deemed responsible for the large size of the primate brain, so that some evolutionary anthropologists, biologists, and colleagues in related fields postulated the hypothesis of a Machiavellian intelligence and the social brain hypothesis. And yet, the concept of social brain is not reducible to the individual manifestations of a social world around someone simply because the brain’s architecture rather reflects forms of social organization, language, and culture. On the other hand, one must actually avoid speaking of “social brain” to evoke the positivist idea that social behavior can be solely explained by brain functioning, as if the brain were the biological substrate that determines sociability and human psychology, without taking into account reflexivity and social, cultural conditionings. The neurobiological and sociocultural evolutionary variables do seem to interact both ways, rendering the task of reconstructing the social brain even more complex and challenging. It is against such a complex semantic context that processes of moral decision-making that materialize in everyday life (instantiated in day-to-day, off-line activities and social interactions) and social media (which instantiate online, particularly in Facebook users and social behavioral games, such as Ultimatum and Dictator) can be measured in neuroimaging experiments. These processes are thus investigated from the standpoint of the neural basis of decision-making, combining both empirical findings and theoretical assumptions, as one of the most intriguing tasks of neuroethics lies on the very level of its normative grounds, namely, whatever accounts for the moral justification of doing the right thing in given circumstances that can be described with the aid of neurotechnologies. The descriptive and experimental dimensions of most experiments fail to provide for such a moral justification, insofar as causality or causation cannot be taken for granted or satisfy ought-like normative claims – after all, not every correlation turns out to be causal.

² MCTI/CNPq N ° 405998/2012-0, MEC/CAPES N ° 18/2012, Researchers: C. Borges, A. Buchweitz, J. da Costa, N. de Oliveira, A. Franco, F. Karasek, N. Madarasz, C. Nahra, F. Pontin, M. Portuguese.

Let me offer in closing a few remarks on the fate of neuroethics in Brazil and elsewhere in Latin America. Just as it happens in other areas of cutting-edge, interdisciplinary research in Latin America, most research in bioethics and neuroethics in this part of the world has been conducted in cooperation with U.S. and European institutions through joint research programs, exchange initiatives, and international events that receive support or intellectual inspiration from major universities and programs in the Northern Hemisphere. With the evolution of neuroscience and neuroethics, models of human social cognition that are grounded in a new range of neuroimaging data also emerged. Given the fast-growing interest in neurotechnologies and neuroscientific research in Latin America, especially in Brazil, Argentina, and Chile, neuroethics will certainly become one of the most important areas of interdisciplinary, cutting-edge research in the next decades. The tremendous potential for human empowerment and social impact brought about by neuroethics attests such an optimistic prognosis, without invoking any utopian ideal of trans-human or post-human scenarios. In effect, the conjugation of the “social brain” with neuroenhancement tends to be rather regarded as part of strategic investments and improvements in public health, so as to make biotechnologies more accessible to larger segments of society. Most brain research centers are thus somewhat committed to this social dimension of public health, as life expectancy and the population of aged people continue to grow in most South American countries and the neuroscientific study of the aging brain, especially the development of Alzheimer’s and Parkinson’s diseases and aging-related dementias, seek to investigate how sensory, motor, sleep, cognitive, and emotional functioning ultimately influence the quality of life of older individuals. Neuroenhancement – even when primarily conceived in terms of cognitive and social-behavioral enhancement for healthy individuals without mental illness—tends to be more and more broadly conceived with a view to improving the processes of aging and minimizing age-related cognitive decline. It is indeed a salient feature of ongoing research in cognitive neuroscience and neuroethics to stress the “social brain” intertwining of emotion, memory, consciousness, and rational decision-making processes in both individual and collective existence. Another feature that remains salient in ongoing research in neuroethics – in Latin American and elsewhere— is that many active groups, centers and researchers are linked to confessional institutions, notably Catholic universities, and this is very important as we consider the impact of neuroscience on worldwide reception of new technologies, particularly those that seem to defy traditional conceptions of human nature. It is thus very interesting to take into account the conjugated processes of democratization and secularization which shaped most Latin American societies, constitutions, and legislations after the several decades of military dictatorship that haunted almost all parts of the subcontinent, following the Cuban Revolution of 1959. Although most Latin American societies can be regarded as postsecular, constitutional democracies, one can observe varying degrees of liberal pluralism and remarkable contrasts between religious institutions that are clearly more secularized and those which remain more conservative and resistant to biotechnological and neurotechnological

innovations, including the spousal of neuroethics. Thus, in order to fare well between Scylla and Charybdis, long-term perspectives for the ongoing, promising research in neuroethics in Brazil and Latin America must systematically avoid these two major extremes: necessitarian thought control and lack of scientific rigor in their research programs.

NATURALIZING INTENTIONALITY

RUTH GARRETT MILLIKAN

Brentano introduced the term "intentionality" into our modern philosophical vocabulary to denote the property which, as he thought, distinguished the mental from all other things. This property is sometimes informally called the "ofness" or "aboutness" of perceptions, thoughts, sentences and so forth. Brentano equated intentionality with the capacity to bear a real relation to something nonexistent, for example, the capacity of a belief to bear a correspondence relation to a nonexistent fact, hence to be false. Similarly, when our intentions are not fulfilled, they seem to bear relations to nonexistent facts. Call this problematic relation "Brentano's relation."

Brentano was surely mistaken, however, in thinking that bearing a relation to something nonexistent marks only the mental. Given any sort of purpose, it might not get fulfilled, hence might exhibit Brentano's relation, and there are many natural purposes, such as the purpose of one's stomach to digest food or the purpose of one's protective eye blink reflex to keep out the sand, that are not mental, nor derived from anything mental. Nor are stomachs and reflexes "of" or "about" anything. A reply might be, I suppose, that natural purposes are "purposes" only in an analogical sense hence "fail to be fulfilled" only in an analogical way. They bear an analogy to things that have been intentionally designed by purposive minds, hence can fail to accomplish the purposes they analogically have. As such they also have only analogical "intentionality". Such a response begs the question, however, for it assumes that natural purposes are not purposes in the full sense exactly because they are not mental. It also fails to explain why this supposedly merely analogical intentionality should be unaccompanied by any analogical "aboutness." Surely one's stomach and one's reflexes are not "about" anything, even analogically. Brentano's relation seems, in these cases, to have come apart from aboutness, another mark usually taken to characterize the mental.

There is another, more positive, way of thinking of natural purposes--one on which a naturalistic theory of intentionality can be built. I will talk about a few of the more interesting features of this response as I have come to understand it.

One thing that the example of natural purposes shows is that although being subject to Brentano's relation may be necessary to intentionality, if we take

¹ Originally published in *Philosophy of Mind, Proceedings of the Twentieth World Congress of Philosophy Volume 9*, Bernard Elevantch, Editor, Philosophy Documentation Center, 2000, pp.83-90 (kindly provided by author).

intentionality to imply aboutness it is not sufficient. Aboutness is associated with a purpose only when the purpose is *explicitly represented*. On the other hand, for there to be an explicit representation of a purpose, *there must first be a purpose* to represent. The naturalist challenge here is to show, first, that the phenomenon of natural purposiveness can fulfill this second requirement, that a natural purpose may, equally naturally, be a represented purpose. Second, it needs to show how the same kind of analysis can be used also to naturalize intentionality in cases where facts are represented rather than purposes or ends. It must be shown, for example, how this way of naturalizing Brentano's relation can apply to the intentionality of beliefs, and of sentences that state facts. It needs to be shown in every case that what creates the appearance of Brentano's relation is merely an underlying *natural* purpose.

The job of articulating the notion of a natural purpose so that it will do what is required of it here, it must be emphasized, is NOT that of analyzing anyone's *conception* of natural purpose, but rather of producing a notion that will organize certain natural phenomena for us in a way that casts light on the apparently paradoxical nature of intentionality. A definition designed to capture such a sense of natural purpose was proposed in (Millikan 1984). I called natural purposes of this sort "proper functions," meaning by "proper" a thing's "own" functions (Latin *proprius*, as in "property"). A thing's proper functions are effects which, in the past, have accounted for selection of its ancestors for reproduction, or accounted for selection of things from which it has been copied, or for selection of ancestors of the mechanisms that produced it according to their own relational proper functions, it being their function to be guided by certain variable aspects of the environment in this production. Whatever has proper functions must have had predecessors that historically effected such functions, thus helping to account for its existence or presence.² Because a thing's proper functions are such relative to its history and do not arise from its current dispositions, it is possible that it may not currently serve, indeed, may not even be capable of serving, all or any of its proper functions.³ This possible gap between a thing's proper functions and what it actually effects or can effect is a naturalized version of Brentano's relation. A thing's proper functions may correspond to states of affairs that never become existent. (There may also be other ways of characterizing natural purposes that will do the same job, but I know of none that has been successful so far.) A great many different kinds of things get reproduced or copied because of the effects that they have. For this reason there are many different kinds of things with proper functions. Genes can have proper functions. Any of the various things that genes have as their functions to produce, either absolutely or relationally (that is, as a function of variable environmental input) also can have proper functions. For example, body organs and any other inherited traits can have proper functions. Inner states, such as the perceptual and cognitive states of organisms, can have proper functions

² This very rough characterization of the notion "proper function" is filled out in detail in the first two chapters of (Millikan 1984).

³ More exactly how this can happen is explained in (Millikan 1984) Chapter 1.

that vary as a function of environmental input to the genetically programmed systems responsible for producing them. Unlearned behaviors can have proper functions that are either variant or invariant with respect to environmental input. There are also very many kinds of proper functions that are not determined through genetic reproduction. Artifacts that are copied from earlier exemplars because these exemplars had certain effects can have proper functions. Behaviors learned by trial and error, hence copied from earlier behaviors, can have proper functions. Behaviors learned by copying others' behaviors can have proper functions. Especially important, conventional behaviors, including production of specific language forms, as well as other kinds of customs, fads, and so forth, can all have proper functions. What these latter functions are may or may not be understood by those who reproduce the relevant behaviors, for unconscious reproduction or reproduction for unconscious reasons is common.

But a proper function, a natural purpose, I have said, only becomes associated with intentionality when the purpose is explicitly represented. The notion of representation that we need here is, I believe, kin to the mathematical notion of representation. According to the mathematical notion, a structure consisting of a set of abstract entities along with certain designated relations among them is said to represent another such structure if it can be mapped onto it one-one. Similarly, an intentional representation corresponds to the affair it represents as one member of a whole set of possible representations. These bear certain relations to one another such that, ideally, the whole structure maps one-one onto a corresponding structure of possible representeds. When it is a natural purpose that is represented, this correspondence relation correlates the representation with a state of affairs that it is its proper function to guide a cooperating mechanism to bring about. "Cooperating mechanisms" are ones that have been selected or tuned to cooperate with one another to perform a certain function or functions. The forms of the representations in the system vary systematically according to the forms of the affairs it is their proper function to bring about, more exactly, to guide cooperating interpreting mechanisms to bring about. The explicitness of these representations of natural purposes results from contrast contrast with alternative purposes that could have been represented instead by contrasting representations in the same representational system.

This explains the intentionality of explicitly represented purposes or goals. Discussing this kind of intentionality before that of representations of facts departs widely from the contemporary tendency which is, often, to ignore the intentionality of explicit purposes and goals completely. Indeed, a common assumption seems to be that the intentionality of single words or concepts can be explained first, next the ways these are combined to express full propositions, and only then an analysis of the how the various sentential moods and propositional attitudes function can be given. On the present analysis there are two fundamental varieties of intentionality, two basic "directions of fit"⁴ for intentional entities, the goal- representing direction and fact-representing direction, and there is no intentionality at all without direction of fit. There is no

⁴ Compare Anscomb (1957).

intentionality without reference to full truth-conditional or satisfaction-conditional content. The intentionality of words and concepts is abstracted from their appearance in more complete functional representations.

The intentionality of representations of fact is not determined by their proper functions. That it represent a fact is a matter of HOW the fact-representation performs whatever functions it happens to have. Every device that has a proper function is backed by a history of devices like it that have actually performed that function, or is produced by a device (concept forming mechanism, belief forming mechanism) with a relational proper function backed by such a history, and so forth. Typically there will be a univocal general causal explanation of how performance of the relevant function was effected in these historical cases, the same explanation for each instance, or close enough. In the case of a fact representation, this explanation makes reference to a certain kind of initial condition. Namely, an aspect of the representation bears a specified mapping relation to a certain condition, typically, a condition in the environment, which relation helps account for the fact that cooperating mechanisms guided by the representation are enabled to perform their proper function or functions under that condition. The mapping is of the sort referred to earlier, correlating a set of possible representations with a set of possible represented conditions, where any such correlated pair would have caused performance of the same proper function(s) in accordance with the same general causal explanation.

This way of naturalizing intentionality has several effects on the analysis of language that are of interest.

Notice first that the analysis applies not merely to the intentionality of perception, thought and conventional language forms, but also to that of other reproduced artifacts such as conventionalized pictures, diagrams, charts, and other notational systems. It also applies to animal signals, such as the dance of the honey bee and the danger thump of the rabbit. Clearly the analysis does not derive the intentionality of thought from that of language. Just as clearly, it does not derive the intentionality of language and other conventional representational forms from that of thought. As the proper function of the nectar-locating dance of a honey bee is not to transmit an intentional attitude from bee to bee,⁴ the proper function of a conventional representation need not be transmission of an intentional attitude either. I have argued, for example, that it is not a proper function of sentences asserting identity, asserting existence, or sentences of the form "x means y" to impart intentional attitudes. Their functions are, rather, to alter the inner representational systems of hearers. These sentences do have truth conditions, however. Their truth conditions concern words. But it is not their function to produce thoughts about words, but rather, to alter the ways in which these words are handled or reacted to by hearers. Can entirely different matter.

This analysis of language agrees with Wittgenstein's insight that language forms, like tools, have jobs to do and that these jobs don't always require them to contribute to truth conditions. On the other hand, it denies that there are any such things as rules of language. Thus it avoids the need to explain the *status* of language rules, so puzzled over in the post-Wittgensteinian literature. Language forms have only proper functions. They have effects that have helped account for

their continued reproduction, for their repeated use, in a given language community, and they have, of course, typical ways of producing those effects. The forces of selection that proliferate a specific language device, such as the English imperative mood form, along with a symbiotic hearer response to it, are comparable to the forces that proliferate mating displays, territory-marking behaviors, danger signals, and so forth, behaviors specific to the various non-human animal species. Of course the underlying mechanisms of reproduction and selection are quite different. But in both cases, the reproduction of a pairing between stereotyped advances by one partner and stereotyped responses by the other depends on the fact that benefits sometimes accrue to both partners. Consider the imperative mood as an example. There must be a high enough proportion of cases in which hearers gain from complying with imperatives, along with enough cases in which speakers also gain by hearer compliance, to keep the form from dying out of the language. Similarly, producing true beliefs in hearers is a proper function of standard indicatives. It is because they have often acquired true beliefs this way that hearers continue to decode the language as they do and, often enough, to believe what they hear. And it is only when hearers believe what they hear that speakers are encouraged to continue to use indicative forms. It does not follow that there is a rule somewhere, or a sanctioned prescription, to the effect that hearers must comply with imperatives. Nor is there a rule that speakers must intend them to comply, or a rule that speakers must speak the truth, and so forth--any more than there are rules or prescriptions for peacock mating displays. Of course, if one is trying to display like a peacock, then one succeed or fail, and if one is trying to speak like an English speaker, one can also succeed or fail.

Proper functions do not concern norms in any evaluative or prescriptive sense. They do not concern norms in a statistical sense either. On the contrary, there are many items that usually fail to perform their proper functions. Although the function of the protective coloring on small animals of many species is to prevent them from being seen by predators hence from being eaten, most members of most small species are eaten anyway. Similarly, consider how few times one has to give in to a child who teases for things to keep the teasing from dying out. It is not because imperatives are usually complied with that causing compliance is one of their proper functions. There are, of course, such things as linguistic norms or standards, but they are such only in the non-evaluative, non-prescriptive and also non-statistical sense in which preventing it from being eaten is a standard or norm for the mouse's protective coloring.

I'll mention one last result of this naturalist analysis of intentionality. When fully spelled out, the description of proper functions implies that every artifact produced for a purpose has that purpose as one of its proper functions.⁵ An utterance produced with a purpose is such an artifact. The purpose or intention with which a speaker utters her words endows these words with a proper function. But if the speaker speaks a public language, then the history of the language forms she uses also endows her words with a proper function. These

⁵ Millikan 1984.

two sources of functions may be consonant, or they may conflict, as in cases of insincerity. When the speaker purpose and the public function of the words fail to coincide, the speaker may or may not intend that the hearer fulfil the speaker's purpose knowingly. In lying, for example, the speaker does not intend this, but in the case of Gricean implicature, the speaker does. One reasonable conception of pragmatics places it where these two sources of function intersect, contrasting it with semantics, which studies the conventional, and studies what I have called the "proper" functions of language forms.⁶

⁶ For a parallel description of language conventions see (Millikan 1998).

MENTAL ELEMENTS AS PART OF A SCIENTIFIC ARCHITECTURE

SOFIA INÊS ALBORNOZ STEIN

1. THE ISSUE

In this article, I want to explore the possibility of seeing private phenomenal experiences along the line with Russell's thought: as having objectivity, which would allow seeing them as objects of scientific research. Or, in other words, to ask if it would be possible to speak scientifically about mental events usually called impressions. Although some philosophers still consider this problem to be inscrutable —the so-called explanatory gap that leads to all the complex and extense debate about qualia, my hypothesis is, as Sellars indicated, that it is highly probable that only errors or limitations in our use of vocabulary impede us in finding a solution according to Russell's intention.

According to McDowell (1996 [1994], 1998, 2009, 2011), sensory impressions are impingements of the world on a possessor of rational capacities. But for him, following the footsteps of Sellars, the empirical content should not be seen only as a kind of "responsiveness to impressions." He maintains, beyond that, that in order to explain the differences between our empirical knowledge and "primitive" imprints, it is not necessary to create a dichotomy between the normative of our rational capacities, which organize empirical material, and the natural. The natural, in turn, should not be viewed as being equivalent to the material, external or corporeal causes, resulting in imprints [or impressions].

How can we maintain empiricism, McDowell questions, given the fact of the different natures of our rational capacities and of our perceptual abilities? The latter seem inseparable from a context governed by physical laws explaining causal networks in a world we call material. And we have no evidence for a specific causal chain between sensory impressions and some of our rational operations. We do have, obviously, and this agrees with Davidson (2001), scientific knowledge about the simultaneity of rational operations and physicochemical events, but this does not suffice as evidence of causal links between them.

A further crucial point of McDowell's argument is his attempt to preserve the empiricist belief in a court of experience, but in its kind not contradictory to the conceptual way of human knowledge, already naturally prone to "be conceptual". In human experience, we would not find anything contradictory to the conceptual and propositional mode of expression of human knowledge; on

the contrary, experience, from its origin, would already be, in the impressions, harmonized with the propositional form of expression of experienced contents. Thus, McDowell's claim that there is no real dichotomy between the normative and the natural seems sound; expressed only in this very simple way, it is also "ecumenical" in the sense that it can be accepted by naturalists from very diverse strains.

Contemporary naturalism – as can be also seen in McDowell's neo-Hegelian lines of thoughts – carries with it the virtue of seeking to pacify the relationship of human sciences with those of natural sciences. However, the dichotomous discourse with which we express the path from impressions to propositions does not favor the explanation of the natural process developing from one to the other.

Moreover, seen in its rationalist bias, modern philosophy is a force initially antagonistic to contemporary naturalism. For in most naturalist tradition, one presumes human beings and all their capacities as part of the natural world; and a transcendental proof is no longer required. So McDowell's question on how we respond to the empirical world shown us by means of the imprint is not, strictly speaking, the same question as Kant's, because it presumes our insertion in the natural world; it does not try to prove it. Of course, McDowell appeals to Hegel in his rejection of the dichotomy of phenomena versus things in themselves; but that is a polite way of saying that our empirical nature is presupposed.

Minimal empiricism proposed by McDowell, which he presents in *Mind and World* (1996 [1994]), follows, as I mentioned, the steps of Sellars, who in *Empiricism and the Philosophy of Mind* (1956) had sought to disprove one of the main theses of classical empiricism, the Myth of the Given; there he showed, however, how to continue admitting the truth of ordinary discourse on impressions and thoughts.

From this theoretical landscape, I want to show why Bertrand Russell's neutral monism already presents naturalistic arguments conciliating third-person causal, scientific and physicalist explanation [on the one hand] with phenomenalist first-person perspective [on the other]. So my argumentative trajectory will not focus exclusively on Sellarsian criticism of the Myth of the Given of empiricist tradition, but will seek to show that there is, in this tradition, a sign of reconciliation between naturalism and phenomenalist first-person perspective. This, perhaps complemented by Sellars' criticism against first person discourse about sensory data, of Wittgensteinian inspiration, and against ontological dogmatic claims, may help investigating human mind, as several philosophers have tried to do.

Keeping the philosophical and scientific discourse about sensations or sensory data by means of its naturalization, of course, as we shall see, does not respond completely to Sellars' criticism, because the emphasis on naturalization does not dissolve the critique of the semantic problem involved in empiricist descriptions of first and third person, Sellars' critical focus on the Myth of the Given. But here we try to safeguard the importance of Russell's neutral monism for the contemporary philosophy of mind reinterpreted by Sellars within a bias of philosophy of language. My analysis aims then to agree with McDowell in his

claim to defend the empiricist tradition, from an epistemological point of view, without, however, and now no longer in line with McDowell, to give up showing the possibility of using this same tradition to support a kind of metaphysics of first person experiences, which, in the philosophies of Sellars and McDowell, are circumscribed by transcendental and semantic frames.¹

The rescue of empiricist metaphysics in philosophy of mind has the potential to bring the latter to the current disciplinary scenario of neuroscience, which is of scientific nature. Limiting the knowledge of the mind to our semantic ability to describe it (cf. Sellars) or to think it (cf. McDowell) is undoubtedly part of the truth about our cognition, because it points to an essential feature of any language, natural or scientific, namely, that any ontology depends on the way we conceptualize. Nevertheless, the realism we presume in the philosophy of mind must not, as I understand it, be a Hegelian rationalistic realism, which establishes a link between objects and concepts, or between facts and thoughts (as in Frege [1986] and McDowell). Realism must be, in a sense, weaker than this: it should assume rational and conceptual limitations of empirical knowledge. Simultaneously, however, it must assume that it is possible, by scientific advances, to achieve greater insight into the human mind, and that is not *only* an ontologically relative attempt to describe mental phenomena, but envisages advances toward more reliable descriptions of human mental processes.

2. RUSSELLIAN NEUTRAL MONISM: ITS GOALS AND CONCEPTUAL TRANSFORMATIONS

We have already a huge amount of literature about the gap between the ordinary point of view on perceptual experience and the scientific point of view on how the perceptual system operates. No discussion to date has resulted in the unification of these perspectives, neither through complete reduction, nor by means of a comprehensive theory that could show the interactions between objects established by both points of view. On the one hand, [we have] the objects of subjective experience mediated by inter-subjective language learning: impressions, sensations, perceptions, etc. And on the other hand, the objects of anatomical and physiological theories: tissues, nerves, electrical impulses etc.

Today, the neurosciences seem to be showing that "subjective" impressions are generated from a causal network that can be traced scientifically. This does not prove that there are sensory data, similar to those described by Russell, but it corroborates the *kind* of causal theory of perception that Russell was working on between 1910 and 1930.

Scientism as drawn up by Russell in the 1920s leads him, for example, to deny the existence of nominal entities such as the "subject":

The subject, however, appears to be a logical fiction, like mathematical points and instants. It is introduced, not because observation reveals it, but because it is linguistically convenient and apparently demanded by grammar. Nominal entities

¹ I presuppose here, in my use of the expression 'metaphysics', that I will be speaking about a phenomenon that is real, but not immediately know by scientific and empirical methods.

of this sort may or may not exist, but there is no good ground for assuming that they do. The functions that they appear to perform can always be performed by classes or series or other logical constructions, consisting of less dubious entities. If we are to avoid a perfectly gratuitous assumption, we must dispense with the subject as one of the actual ingredients of the world. But when we do this, the possibility of distinguishing the sensation from the sense-datum vanishes; at least I see no way of preserving the distinction. Accordingly the sensation that we have when we see a patch of colour simply *is* that patch of colour, an actual constituent of the physical world, and part of what physics is concerned with. (Russell, 1921, p. 142)

For our inquiry it is important to ask why Russell eliminates the subject, when he assumes the perspective of neutral monism, in 1921, while maintaining the sense-data, now equivalent to sensations, neither purely mental nor purely physical? We can venture an answer: because to him it seems right and necessary, based on the principle of epistemological atomism in which he operates, that sensations are something real and, in a sense, physical. Without this assumption, the causal link between objects and perception would be jeopardized. But to do so, he does not think it necessary to state that sensations are determined, in their content and nature, by a relationship with an intentional and conscious subject.² As Hatfield explains:

As developed by James and Russell, neutral monism avoided the mind-body problem by positing only one “stuff”, the allegedly neutral “stuff” of momentary particulars, or pure experiences, or Machian elements. Mach, James, and Russell could then point to two sets of laws to be found empirically in the successive states of this stuff: psychological laws governing successions of perceptions and other mental states considered as mental, and physical laws governing successions of perceptions and posited sensibilia, or unexperienced pure experiences, considered as physical. The mind-body relation then became a matter of tracing connections between physical sequences and intersecting psychological sequences of momentary particulars (as in Russell, 1921, ch. 15). (Hatfield, 2002, p.222)

Thus, despite the differences in the explanatory laws of the exclusively physical and the so called psychological, the momentary individuals, or elements, in Mach vocabulary, need not be called strictly physical or psychological. They would be "neutral" and could be explained by laws both physicalist and psychological.³

² I will not deal here with the many questions that this statement raises, for example, if there is a unity of mental representations, if there is a point of convergence of representations, which would be a consciousness without a subject or an "I", as the mind could be intentional without the assumption of an "I", who would be the agent of intent, etc.

³ Hatfield, in his analysis, maintains an ontological dualism when he says that the discourse about the neutral stuff (neutral elements) could be both a discourse about psychological or physical states.

Today, the neurosciences do not prove the existence of sense data, nor of sensations that would be both the sense experience and the sensory data themselves –depending on the perspective from which these were explained; yet [neurosciences] preserve Russell's aim of finding the causal net ranging from physical objects to their "mental" perception. And that network must include various forms of "representations" of object properties. The observation of neuronal activation in the process of perceiving physical objects led to the conclusion that perception goes through several stages and that, in a broad sense, the objects really are gradually built up in our brain, through the inter-relationships being established between different parts of it, each performing a different function.

Obviously the access to one's own perceptions and sensations is still of first-person only, which shows that Russell's statement (1921) is still quite present [which says] that sensations are the intersection between mental (first person) and physical (third-person). It is also obvious that the search for the causal network bringing objects to their perception does not exempt from inquiring about what we call "representation", a concept still extremely controversial both in philosophy and neuroscience. And once this is the case, methods and testing must take into account the first person report on representational "subjective" experiences.

We are now facing a widespread dilemma in the scientific community: we *either* accept the common *descriptions* made by humans about their own subjective experiences (their "talk about impressions") and correlate them to physical changes we can scientifically observe, like parallel events, *or* we try to investigate, in a behaviorist approach, what is happening physically and in terms of their behavior when they are in a given position of perception, and discard the first-person mental description, linking initial physical behavior to resulting physical behavior. In any case, the link between mental or psychological experiences and starting or ending physical behavior is yet to be discovered. This link Russell thought he had found in the "sense data", redefined as "sensations" in 1921.

According to Savage (1989), the terminological change wrought by Russell in 1921 could really be considered as redefining the concept of sense-data. What some considered to be an abandonment of the sense-data concept as such, is seen by Savage as abandoning the allocation of certain sense-data properties:

...we may say that before abandonment sense data were held to be the absolutely certain (indubitable, infallible), immediate (uninferred, self-evident), precise (analyzed, simple) data of empirical knowledge; and after abandonment were held to be only relatively certain, immediate, and precise, i.e. certain, immediate, and precise to some degree. (Savage, 1989, pp. 138-139)

Like Savage, Tully (1988) believes there is continuity from Russell's thought of 2014 until the publication of "The Analysis of Matter" (1927), through "The Analysis of Mind" (1921). And that therefore terminological changes were not radical conceptual changes, but were redefinitions of concepts already used from the beginning. Initially, in 1914, Russell is still critical of the new realists, like

William James and Ernst Mach, considering their positions a throwback, a rapprochement to an idealism of sensations, according to which sensations were reduced to ideas, which thus would become the only element of reality, an idealism that he himself tried to overcome by the notion of familiarity ("acquaintance") between the subject and the data coming from the senses. However, as Tully points out (1988), contrary to Russell's criticism, the New Realists did seem aware of the challenge to find exactly those entities that could be considered both mental (psychological) and physical (or, maybe, material) as well:

Interestingly, the New Realists appear to have been aware of the problem and, officially at least, designated neutral items as "qualities"; but he paid no attention to this when he accused their theory of being tainted with idealism. Whether a word like "quality" would in fact achieve what is wanted may be questioned, just as it may be wondered whether Russell's own use of the expression "sense-datum" could ensure the degree of objectivity and independence he required for the construction of physical objects. The problem of basic descriptions is a common one for empiricists, not just for those who defend neutral monism. (Tully, 1988, p. 214)

Thus it is not surprising to his interpreters that Russell, at the end of the 1910s, has begun to find attractive the notion of "neutral element", coined by Mach, an element neither merely psychological nor merely physical, strictly speaking.⁴

From 1921 Russell began to use the word "sensation" instead of "sense-datum", but this terminological shift did not mean exactly a withdrawal from certain russellian empiricist philosophical trends, which had as one of its lines of thought the search for an empirical basis for explaining mental phenomena. This interpretation of Russell's philosophical development in the philosophy of mind is partly justified because he changed terminology on several occasions. It was not, according to Tully (1988), the neutral monism that led him to choose the expression "sensation" to replace "sense-datum", but apparently he was looking

⁴ According to Banks (2010), based on his analysis of the discussions of the first decades of the twentieth century, the three versions of neutral monism, William James', Ernst Mach's and Bertrand Russell's, can be roughly summarized by four theses: "This historical climate of ideas ... left its stamp on the following theses about the nature of elements. I do not say that these theses are articulated by every classic neutral monist, but they are implicit in all three of them:

1. Monism: the mental and physical domains are part of a greater natural domain of elements and their functional variations.
2. Neutralism: elements are neither mental nor physical; rather minds and physical bodies are functionally related complexes of elements. Certain functional variations of the elements are called 'mental' and others are called 'physical', but there is no underlying duality of variations.
3. Psychophysical Identity: every sensation, such as green, is also a physical element (s/e), a neural energy in the brain. Not every element is a sensation, or even a possible sensation.
4. Powers: elements are powers with causal force. They are concrete qualities and dispositional ways of affecting things in their various causal, or functional, roles. The concrete quality instantiates the dispositional, relational role. Every element is naturally embedded in its functional role." (Banks, 2010, p.177)

for the most appropriate term to describe something not strictly speaking psychological, and which should be possible to describe from a science of the human mind. Soon after employing the word "sensation" replacing "sense-datum", in 1921, Russell replaced it by the words "percept" ["percepto"] and "quality". The decreasing emphasis on the introspective method and on the concept of *acquaintance* in writings prior to "The Analysis of Mind" (1921), like in "The Relation of Sense-data to Physics" (1914b) and "On Propositions: What They Are and How They Mean" (1919), seems to have contributed to Russell's becoming closer to the understandings of the new realists of mind, such as James and Mach. Russell says in "On Propositions" that James' neutral monism was making more sense, as he was realizing the importance of linking the physical objects of scientific research, particularly the brain, to the investigation of mental activities. According to Tully:

It is clear that Russell thought of neutral monism as a corrective to the other two varieties of monism which suffer from quite different but equally severe limitations, the one because it refuses to recognize the powerful role of inferential knowledge in science, the other because it fails to provide a systematic account of the ultimate dependency which such knowledge has on first-person experience. (Tully, 1988, p. 222)

Tully (1993) tries to explain Russell's shift in 1914 towards neutral monism, after criticisms he made in the same year (1914a) to the new realists – which includes his criticism on belief as formulated by them, and on the absence of a satisfactory explanation of the first-person (subjective) experience; this shows the importance given by Russell to the unification of the discourses of physics and psychology, which until 1914 were considered irreconcilable:

In treating sensory particulars as the common subject matter of both physics and psychology, Russell was adopting one of the tenets of the New Realism (or neutral monism, as the doctrine soon came to be called). In fact, the list of criticisms of neutral monism which he published in 1914 omitted this tenet; what he objected to rather was the neutral monists' account of belief and their tendency to dismiss the importance of first-person experience. Unobjectionable on other grounds, the hypothesis must have satisfied his desire for logical economy, but if Russell wanted to classify sense-data (or sensibilia) as physical, whatever qualitative familiarity they were allowed to have, how did he think these particulars become part of first-person experience? What kind of relation occurs when a sense-datum is experienced? Evidently, it takes more than a sense-datum plus the presence of a human body with functioning sense organs and a brain, for such things are themselves physical constructs whose ultimate constituents (in the eyes of physics) would mostly be sensibilia (or "ideal" appearances). The existence of a brain might be necessary for the experience of sense-data but is hardly sufficient. To this question Russell offered a forthright answer: "If —*per impossibile*— there were a complete human body with no mind inside it, all those *sensibilia* would exist, in relation to that body, which would be sense-data if there were a mind in the body. What the mind adds to *sensibilia*, in fact, is merely awareness: everything else is physical or physiological" ("The Relation of Sense-data to Physics", p. 8). (Tully, 1993, p. 12-13)

From 1914 to 1921, when he published "The Analysis of Mind", Russell reworks his ideas about the sense-data, adapting them to the positions of the new realists. But despite the conceptual and terminological changes, and the substitution of "sense-data" by "sensations" in 1921, as Tully says, this substitution is less substantial than it appears: "Sensations have now been installed in place of sense-data, as required by the new thinking, though ironically their status (as particulars) was virtually identical to what had been previously accorded to sense-data." (Tully, 1993, p. 34)

Also Gary Hatfield (2002) warns against a hasty understanding of Russell's position in 1921, prompted by claims of Russell himself in "The Analysis of Mind", thus agreeing with Savage and Tully. To justify the terminology replacement operated by Russell, [Hatfield] links this with the aforementioned refusal of maintaining a subject to which the sense-data would be presented. The subject would be unnecessary in the case of "sensations", which would be sequences of momentary particulars, not needing to be particular "to a subject":

When Russell became a neutral monist he also came to reject sense-data (1919/1956, p. 306; 1921, pp. 141-2; 1959, p. 135). Care must be taken in interpreting this change. In rejecting sense-data, he did not reject sequences of momentary particulars. Rather, he came to reject the distinction between such particulars and a subject who senses them. The particulars were no longer to be regarded as "data" for a subject, because the subject itself was denied. Or, to put it another way, the experience had by a certain subject is now regarded simply as a specific sequence from among the various sequences of momentary particulars that constitute everything, and which are the only particulars whose existence is explicitly allowed. (As we shall see, Russell did not flatly deny that the subject exists. but he took the theoretical attitude that its existence was not needed and should not be posited.) Russell's immediate particulars are now to be equated with the "elements" of James and Mach (as noted in Russell, 1919/1956, p. 305; 1914a/1956, p. 140). (Hatfield, 2002, p.208)

Quine (1966) condemns Russell's (1921) maintaining of a speech on "neutral particulars" which, in his point of view, resembles the talk about sense-data from previous works, but praises the strengthening of naturalism in the fact that Russell assumes neutral monism:

Russell meanwhile was warping his logical atomism over from its frankly phenomenalistic form to what, influenced by Perry and Holt, he called "neutral monism" (Cf. *Analysis of Mind*, p. 25; *Analysis of Matter*, ch. 37). Neutrality here has a bias, as it often has in politics; Russell's neutral particulars are on the side of sense data. Still, a drift has begun, and it continues. It does not reach the physicalistic pole, even in *Human Knowledge*; but there is an increasing naturalism, an increasing readiness to see philosophy as natural science trained upon itself and permitted free use of scientific findings. Russell had stated the basis for such an attitude already in 1914: "There is not any superfine brand of knowledge, obtainable by the philosopher, which can give us a standpoint from which to criticize the whole of the knowledge of daily life. The most that can be done is to examine and purify our common knowledge by an internal scrutiny,

assuming the canons by which it has been obtained" (*Our Knowledge of External World*, p.71). (Quine, 1966, p. 667)

Neutral monism, in the hands of Russell, gained new empiricist and physicalist contours, benefiting from years of theoretical experience of one of the masters of logic and epistemology of the twentieth century. His determination to avoid idealism was not enough to take him closer to a more extreme naturalist position as coveted by Quine, but his choice of neutral monism avoided his adherence to behaviorism and eliminativism, which hampered Quinean philosophy getting closer to neuroscience at the end of the last century.

3. RUSSELL'S THEORY OF PERCEPTION: OSCILLATING BETWEEN PHENOMENALISM AND PHYSICALISM

In "The Analysis of Mind" (1921), Russell says that "we perceive things more or less, but always with a very considerable amount of vagueness and confusion" (p. 135). From his theory of perception is not possible to derive a strong realist position about perceived physical objects. Russell oscillates between the classical empiricist position and physicalism. For him, if we can explain the causal relationship between the object, the environment and the human body, it is possible to establish whether the perception is true or not. Russell says:

When a mental occurrence can be regarded as an appearance of an object external to the brain, however irregular, or even as a confused appearance of several such objects, then we may regard it as having for its stimulus the object or objects in question, or their appearances at the sense-organ concerned. When ... a mental occurrence has not sufficient connection with objects external to the brain to be regarded as an appearance of such objects, then its physical causation (if any) will have to be sought in the brain. In the former case it can be called a perception; in the latter it cannot be so called. (Russell, 1921, p. 136)

Russell's description of perception events shows the oscillation between empiricism and physicalism in his philosophy of mind of 1921. And besides, we could say, it shows that Russell already espoused a sort of physicalist, non-idealist or solipsist naturalism, because although speaking of appearances and the obscurity of perception content as appearances, he did not hesitate to affirm the sovereignty of physical science to establish the causal links between objects and phenomenal experiences.

In Russell there is not a mere parallel between physical and mental. Appearances should be explainable by physical laws. However, despite this physicalism in his perception theory, there remains a gap between the causative process of perception and the form of appearances. And that is why, as we have seen above, Russell's neutral monism puts the sensations at the heart of reflection on the contents of empirical knowledge:

Sensations are what is common to the mental and physical worlds; they may be defined as the intersection of mind and matter. This is by no means a new view; it is advocated, not only by the American authors I have mentioned (ex: William James), but by Mach in his *Analysis of Sensations*, which was published in 1886. ... It [sensation] is not itself knowledge, but it supplies the data for our knowledge of the physical world, including our own bodies. (Russell, 1921, p. 144)

Russell does not start the discourse on the contents of human mind from the perspective of the first person only, gradually unfolding in an explanation of correlation between contents of the mind and what they designate or should designate, as in the Carnapian attempt (1928) to demonstrate the correlation between the human world as experienced, and scientific-conceptual knowledge. The classical tradition, rationalist as much as empiricist, presupposes certainty about the first person's perspective and "builds" the world and / or the knowledge about it from that perspective. This always happens with the help of cognitive contents that this perspective makes available, that is: when they explain how we reach knowledge from phenomenal content, they presume that the knowledge resulting from this process is already known. Thus, Carnap, for example, shows that it is possible to explain how we feel and perceive already presupposing knowledge we have about the physical process of sensation and perception. For Carnap therefore philosophical explanation should not just start from facts described from a third person, but should show how the process evolves internally in the person who is feeling and sensing. However, there is obviously circularity in the expositive method that no anti-naturalist statement can avoid. So the accusation that could be made to the neo-empiricists, of falling into an idealist or solipsist position for starting the description of the contents of the human mind from a first person point of view does not hold if we bear in mind that every description of first person uses knowledge we call the third person knowledge of the relationship between subject and environment, which only science provides.

This means that the analysis of first person discourse depends on the assumption of third-person knowledge. This in turn means that the discussion about the sensations, participating in perceptions of objects, is a third person discussion which *looks like* a first person discussion.

4. SCIENCE OF MIND AND ORDINARY DISCOURSE: MICRO- AND MACRO-THEORIES

To a certain extent, Sellars also demonstrates the first-person / third-person relationship, but only to the point of denying that the discussion of sensations is useful in explaining empirical knowledge, whilst identifying fallacies in the empiricist way of upholding the Myth of the Given. This is partly because the empiricist affirms the existence of particular entities, sense-data, which are, in fact, the result of a *theory* concerning perception. Sensory data cannot be detected either from the perspective of the first or of the third person. They are therefore arbitrary stipulations (or postulations) of empiricist philosophy, and are treated by the empiricist as if they themselves constituted a form of evidence.

Sellars makes use of the Myth of Jones to explain how the Myth of the Given originated, and how the theory concerning sense-data derives from the attempt to explain the logic of ordinary language when it speaks about immediate visual experiences (Sellars 1997 [1956]: 109):

From this standpoint it is sufficient to suppose that the hero of my myth postulates a class of inner —theoretical— episodes which he calls, say, *impressions*, and which are the end results of the impingement of physical objects and processes on various parts of the body, and, in particular, to follow up the specific form in which I have posed our problem, the eye. (Sellars, 1997 [1956]: 109)

Although Sellars' criticism to the Myth of the Given is well-founded and can't be disregarded, I wouldn't say that it is a categorical illusion to claim that sensory data are first and foremost the content of empirical knowledge, *as long as* third-person knowledge could prove (direct or indirectly) that sense-data do, in fact, exist. In this context, it is interesting to consider the following passage from Sellars:

...some philosophers have thought it obvious that we can expect that in the development of science it will become reasonable to identify *all* the concepts of behavior theory with definable terms in neurophysiological theory, and these, in turn, with definable terms in theoretical physics. It is important to realize that the second step of this prediction, at least, is either a *truism* or a *mistake*. [...]

To ask how *impressions* fit together with *electromagnetic fields*, for example, is to ask a mistaken question. It is to mix the framework of *molar* behavior theory with the framework of the *micro*-theory of physical objects. The proper question is, rather, 'What would correspond in a *micro*-theory of sentient organisms to *molar* concepts pertaining to impressions?' And it is, I believe, in answer to this question that one would come upon the *particulars* which sense-datum theorists profess to find (by analysis) in the common-sense universe of discourse (cf. Section 23). (Sellars 1997 [1956]: 113-114)

We can see from the above quotation that Sellars does not reject a description of sense-data which is part of a scientific theory concerning human perceptions. "Sense-data" may be part of the technical vocabulary used to describe how sentient organisms react internally to external forces. However, this does not mean that the sensory data provide the specific reference for expressions such as "red" or "triangle". Impressions of redness or of a triangular form may be part of a micro-theory, which can explain the reaction of sentient organisms to the environment, even though it would be fallacious to deduce that ordinary language is incorrect to attribute properties to external objects which are inherent to the impressions resulting from our sentient relationship with these objects. So, it would be correct to conclude that the micro-theory *can* explain the relationship between physical objects and the perceptive qualities they induce in us, and that, *in line with the micro-theory*, common sense *could* be wrong to attribute to objects any qualities which may belong to perception. *However*, this

should not lead us to "correct" ordinary language, since in this case we would merely be substituting one theory for another, or substituting the physicalist argument concerning the qualities of external objects for a third-person phenomenalist argument concerning perceived qualities attributed to the external objects which cause them.

If we interpret the philosophy of perception propounded by Russell in 1921 as a kind of scientificism which does not completely reduce sensations to physical events, but which allows us to think of sensations as having a phenomenal or mental aspect, which would be still part of the world that is the object of the sciences, then Sellars' criticism of the Myth of the Given does not run entirely counter to Russell's perspective. This perspective does not use only introspection as a methodology, but relies, rather, on evidence from physical science to explain perception. Nevertheless, Russell could still be criticized by Sellars for giving pride of place to the physicalist perspective, to the detriment of the perspective of common sense, and for not seeing that both perspectives are essentially just two different languages and theories with different pragmatic goals, where one is no truer than the other.

If we analyze Sellars' claims about the differences between ordinary language and scientific language, we can conclude that what happens in both cases is that each of them must refer to distinct objects. *Ordinary language* must refer to physical objects as both its aim and its function; *scientific language* (which investigates the process resulting in the attribution of sensible qualities to physical objects) focuses on describing the process of perception, and not on the intersubjective process of speaking about perceived physical objects. Scientific language wishes to explain the process of perception, even though this process may not be clear to the subject who perceives, whilst ordinary language allows the sentient subject to communicate what is happening in the physical world to her interlocutors, including, in some cases, what is happening in her body or in her mind. The fact that in ordinary language the subject attributes to objects qualities which appear to her in a unique form does not, obviously, mean that she knows which process led to her singular perception of these qualities, or to the attribution of specific words for describing them. Learning how to associate words with specific experiences does not require the ratification of a theory of sensory impressions, which a naturalized philosophy of perception would seek to conceive with the assistance of the sciences. From a pragmatic perspective, the really central question seems to be in which sense scientific explanations of the process of perception (or of other psychophysical processes) are relevant for everyday intersubjective discourse and practices. It is pretty obvious that from a scientific point of view knowing about the psychophysical process is essential for establishing psychological or physical treatments, but it isn't obvious that a scientific discourse could be useful to everyday discursive practices, as a reductionist eliminativism in philosophy of mind could advocate.

SOME METAETHICAL PROBLEMS FOR PRAGMATIC NATURALISM

SUSANA NUCCETELLI

This essay looks closer at Philip Kitcher's 2011 view of the so-called ethical project, which creatively combines elements of both American pragmatism and philosophical naturalism to make up the distinctive metaethical perspective he dubs 'pragmatic naturalism'. As a normative doctrine, pragmatic naturalism proposes an egalitarian conception of the good and a method for ethical decisionmaking meant to be of help in resolving some contemporary moral dilemmas arising from current resource scarcity and global inequalities. As a metaethical doctrine, pragmatic naturalism is an anti-realist doctrine that Kitcher believes can avoid the skeptical consequences affecting other forms of anti-realism. Here I will first look closely at the genealogy of morality that pragmatic naturalism was devised to accommodate. Then, I will explain why pragmatic naturalism falls short of meeting some metaethical challenges facing it.

I. KITCHER'S MORAL GENEALOGY

Pragmatic naturalism amounts to dynamic view of folk morality which assumes an 'ethical-project thesis.' Given this thesis, morality, rather than a static result of natural selection, is an evolving social enterprise. It is an ongoing social project that has been in the making ever since an "ability to apprehend and obey commands changed the preferences and intentions of some ancestral hominids, leading them to act in greater harmony with their fellows and thus creating a more smoothly cooperative society."¹ Because social life always presents new difficulties, the ethical project is *per force* in constant development and therefore unfinished. The main reason for this thesis is its consistence with a plausible reconstruction of the origin and evolution of human morality. The pragmatic naturalist finds evidence from evolutionary psychology and behavioral studies of the Great Apes for the existence of a rudimentary form of normative guidance among our hominid ancestors, which arose in response to their need to work out how to live together. Normative guidance later took stable roots among early humans, who expanded and refined it in response to their social needs. Since the ethical project relies heavily on these hypotheses about the origin and evolution

¹ Kitcher, *Ethical Project*, 74. Further references to Kitcher's book will be given in parentheses in the body of the text.

of morality concerning the role of natural and, principally, cultural selection in its emergence, let's first have a closer look at them.

Appealing to historical facts and plausible conjectures, Kitcher articulates a credible evolutionary story of morality that is generally consistent with other genealogies of morality currently on offer. Yet unlike some of these accounts, his story emphasizes the evolutionary role the *social* environment, something often neglected in narratives focused exclusively on natural selection. In previous work, Kitcher has sought to avoid telling a story that confuses morality with “nice behavior” or relies on the existence of an alleged moral instinct.² According to his historical narrative, normative guidance first appeared among our hominid ancestors as a result of social pressures to regulate their conduct in order to improve social cohesion. Organized in small, mixed groups, these hominids often faced injury, disease, and death caused by extreme scarcity of resources and their own tendency to engage in violent behavior toward each other. But the capacity for normative guidance then evolved through amplification and refinement of norms in response to the increasing demands of their larger societies, and this capacity was later passed on to early humans. Here Kitcher's narrative helps itself to an evolutionary hypothesis commonly found in the literature, which links the emergence of morality to the possession by the original moral agents of a strong, stable (i.e., not episodic) tendency to engage in altruistic behavior. That tendency is taken to have led, through natural and cultural selection, to the development of rudimentary normative guidance. In Kitcher's version of this hypothesis, groups of non-human ancestors in which altruistic behavior had a wider scope (i.e., appeared in more species) and range (i.e., manifested itself in more types of activity) were better than other groups at maintaining the social cohesion needed for responding more efficiently to resource scarcity and violence. Groups with more of such individuals then did better than other hominids from an evolutionary standpoint. It is also likely that the hominids with altruistic tendencies were better disposed to be motivated by rudimentary norms of conduct and to have some co-evolved reactive attitudes (e.g., guilt and resentment) than those who lacked those tendencies.

So, normative guidance was a feature that first appeared among our hominids ancestors, who introduced it to boost social cooperation and punish altruism-failures. These hominids' altruistic capacity is likely to have been inherited by our early human ancestors. In their case, Kitcher's story has it that life in larger groups required more sophisticated normative guidance that could serve to prevent and punish altruism-failures in a wider and more severe way (2011: 74-96). Cultural selection then played a major role in the human expansion and refinement of normative guidance – and thus also in getting the ethical project off the ground.

A crucial hypothesis in this historical narrative is that the original function of ethics was remedying altruism-failures. Support for it is offered by analogy and

² In “Biology and Ethics,” Kitcher explicitly objects to appeals to a moral instinct (e.g., Hauser, *Moral Minds*). For a recent discussion of the analogy between innate language and morality, see Harman, “Naturalism in Moral Philosophy.”

evidence. The fossil record attests to our hominid ancestors' having had a social arrangement of small groups, of the sort explained above. To support his reconstruction of the challenges they must have faced, Kitcher invokes primatologists' observations of chimpanzees and bonobos, whose social life is regularly strained by similar violence and scarcity of resources. These animals too have a tendency to altruistic behavior, suggests Kitcher, which they exhibit toward their "friends" (i.e., other members of their coalitions and sub-coalitions). But since those tendencies are weak, episodic, and often overridden by the animals' strong tendencies to aggression, their group's cohesion is commonly undermined by eruptions of violence. As may be observed in the behavior of chimps – whose grooming activities extend far beyond the animals' hygienic needs (up to six hours a day) – restoration of peace after a period of violence is a time-consuming, energy inefficient process. By comparison, the rudimentary normative guidance of our primate ancestors represents a considerable improvement in efficiency at the remedying of altruism failures.

Kitcher also appeals to game-theoretic models that have been instrumental in evolutionary explanations of observed altruistic behaviors among some non-human animals. According to those models, cooperation among parties comes out as the strategy that maximizes outcome. This has been taken to explain the evolutionary advantage of certain types of altruism – specially, the seemingly altruistic behavior of some animals toward the non-kin. Now called 'reciprocal altruism,' altruistic behavior of that sort has been claimed to be compatible with evolutionary theory by construing it as a variation of 'I rub your back if you rub mine' – a strategy that pays for the giver in the long run, whether or not a particular receiver of the altruistic behavior returns the favor. On the other hand, the compatibility of kin altruism with evolutionary theory was made clear once technology showed the genetic material shared by givers and receivers of this kind of altruistic behavior (observed especially among bees, wasps, and other insects). But neither of these types of biological altruism can account for the emergence of normative guidance and the beginning of the ethical project, both of which require beings with a capacity to be motivated by moral norms and related psychological attitudes. What is needed is psychological altruism, shown by organisms capable of adjusting their psychological states and feelings so that they become more favorable to others' fulfilling of their own wants, plans and intentions. Sensitive to the desires, plans, and intentions of others, the psychological altruist modifies her own psychological states so that they can contribute to the satisfaction of desires, plans, and intentions of others – and this is the only factor that explains that modification.

Kitcher holds that our pre-ethical ancestors had altruistic tendencies of precisely this sort. Those tendencies predisposed them to work out agreements on rudimentary norms of conduct. "Very probably," hypothesizes Kitcher, "they began with precepts about sharing scant resources and not initiating violence" (6). The ability to regulate conduct with precepts, inherited by early humans from their primate ancestors, made possible the ethical project at its initial stages, long before revisions were introduced to accommodate novel challenges to social life. Arguably, the proto-ethical project of our remote ancestors had

scant core norms concerning what we would now regard as distributive justice and cooperation, whose function was the amelioration of altruism-failures within their group, especially the failures that resulted in outbreaks of violence. Ethics had, then, the instrumental function of responding to a certain “problem background,” which relates to Kitcher’s main historical hypotheses as follows:

- (1) The problem background consists in social instability and conflict caused by altruism failure.
- (2) The original function of ethics is to promote social harmony through the remedying of altruism failures.
- (3) Our ethically pioneering ancestors had only a dim appreciation of the problem background, responding to the difficulties and discomforts of a tense and fragile social life.
- (4) We know more about the problem background than they did and offer partial and incomplete diagnoses of the types of altruism failures to be remedied....
- (5) Even with respect to the original function, the project of refining the codes we have continues (2011: 225).

For Kitcher, the positive normative theory consistent with these hypotheses that should be preferred is ‘dynamic consequentialism,’ a view that stands in contrast with traditional deontology and classical utilitarianism. Unlike Kantian deontology, dynamic consequentialism avoids the mistake of prescribing conduct according to rules not well adapted to producing good outcomes, which he finds irresponsible on the basis of familiar counterexamples to the Kantian rule against lying (2011: 288). On the other hand, he criticizes classical utilitarianism for being a static theory of the good and for arriving at its conception of the good by a series of reductionist moves. As a result, classical utilitarianism mistakenly holds that “the value of a world will always consist in the sum of the value of the lives of the individuals we consider one by one” (2011: 292). Kitcher finds this claim questionable on the grounds that it fails to take into account distribution and fails to consider the moral relevance of past events, each of which – argues Kitcher – matters morally. To show how dynamic consequentialism accounts for some issues of applied normative ethics, Kitcher focuses on hard cases, such as accounting for the wrongness of cruelty to animals, or the problem of how to respond to Derek Parfit’s Repugnant-Conclusion puzzle for consequentialism. In addition, he gives details of a rational procedure for moral deliberation (2011: 334 *ff.*), which builds on the reflective equilibrium method developed by John Rawls (1971).

II. SOME METAETHICAL CONCERNS

For Kitcher, his pragmatic naturalism is the metaethical doctrine that best accommodates this moral genealogy and normative theory. On his view, pragmatic naturalism avoids all major problems faced by both moral skepticism and moral realism, including those raised by G. E. Moore’s open question argument and David Hume’s Is-Ought gap. According to pragmatic naturalism, moral norms are entirely explicable in terms of natural and cultural selection. But if so, unless pragmatic naturalism can invoke a suitable conception of moral truth, it would seem committed to a conclusion it is thought to avoid: namely,

that morality is nothing more than a series of changes in normative guidance, with no changes amounting to improvements in any robust sense of ‘moral progress.’ Furthermore, in spite of its anti-skeptical promises, absent a suitable conception of moral truth, pragmatic naturalism would seem unable to account for the authority moral judgment. In fact, pragmatic naturalism would appear to provide grounds for a qualified evolutionary debunking of morality. The argument to show this is primarily epistemic: it starts out by noting that our moral beliefs result from forces, such as natural and cultural selection, that are irrelevant to their truth. It then concludes that there are no reasons for the justification of moral beliefs beyond their contribution to social cohesion. Clearly, since a moral belief might be conducive to social cohesion whether it is true or false, being so conducive may count as a reason for having it, but not as an epistemic reason. If this is correct, pragmatic naturalism provides the grounds for a quasi-debunking (i.e., qualified) argument for morality.

Compare a straightforward debunking argument, offered by skeptics who think that normative guidance can be explained completely as a result of evolution by natural or cultural selection in ways irrelevant to their truth. It typically runs this way:³

1. Given the theory of evolution, our moral beliefs have arisen in ways irrelevant to their being true.
2. If (1), then our moral beliefs are not justified.
3. *Therefore*, our moral beliefs are not justified.
4. *Thus*, theory of evolution debunks morality.

Premise 1 invokes the role of natural and cultural selection in shaping our moral beliefs, a claim that, under some interpretation, no one wants to deny. The plausibility of the conditional in premise 2 depends upon a number of epistemic assumptions. First, its antecedent would imply its consequent, provided there are no reasons independent of moral beliefs’ etiology to think that at least some moral beliefs are true. Second, for the implication to obtain, either the epistemic externalist’s claim (i), or the epistemic internalist’s claim (ii) must be true:

There is no reason to think that evolution (by natural and/or cultural selection) has endowed us with a reliable moral belief-forming mechanism.

If you know that your moral beliefs have been produced in a ways irrelevant to their truth, then unless you have independent reasons for taking them to be true, that knowledge would defeat their justification.

Although a defense of at least one of these assumptions would be needed for premise 2 of the debunking argument to be problem-free, I’ll ignore this complication henceforth, granting that either (i) or (ii) is acceptable. In addition, note that I am casting the evolutionary debunking argument in terms of *moral*

³ For example, Richard Joyce, in *The Evolution of Morality*, 171-219, offers an evolutionary debunking argument along these lines.

belief, although this concept is absent in the account offered by the pragmatic naturalist, who might have omitted it to avoid begging any questions or to show its affinity with certain versions of moral skepticism. Be that as it may, he does invoke instead the ‘descriptive counterparts of norms’ in a context where truth bearers associated with moral norms are at issue. Since this talk is unnecessarily cumbersome, however, hereafter I’ll take truth bearers to be the content of moral beliefs, construing these broadly as to refer to whatever psychological attitude underlines the acceptance of the content of a moral norm. I now want to suggest that pragmatic naturalism fuels the following parallel, quasi-debunking argument:

1. Given pragmatic naturalism, our moral beliefs have arisen to fulfill the original function of ethics (i.e., the remedying of altruism failures).
2. If our moral beliefs have arisen to fulfill the original function of ethics, then they have arisen in ways irrelevant to their being true.
3. If (1) and (2), then our moral beliefs may be justified in terms of their role at fulfilling the original function of ethics, but they are not epistemically justified.
4. *Therefore*, our moral beliefs may be justified in terms of their role at fulfilling the original function of ethics, but they are not epistemically justified
5. *Thus*, pragmatic naturalism quasi-debunks morality.

Since the argument is valid, whether pragmatic naturalism is committed to endorsing it depends on whether it has the resources to doubt its premises – which, as I’ll now show, is by no means clear. First, given the pragmatic naturalist’s moral genealogy and his rejection of moral truth in the realist, representational sense, he cannot make the case that some of our moral beliefs were selected in ways relevant to their being true independently of their role at fulfilling the original function of ethics. The sense of ‘truth’ that creates problems for the pragmatic naturalist here is that invoked by moral realists. After all, as argued by Sharon Street (2006) all scientific evidence points to the conclusion that our moral beliefs are the result of evolution by natural selection, whether they are true or not. But the pragmatic naturalist maintains that some moral beliefs *are* true in a more idiosyncratic, non-representational sense that’s dependent on a prior notion of moral progress. Here is his strategy:

[W]e can make enough sense of ethical progress, prior to any conception of truth, to combat the thought that the evolution of ethics is a history of mere changes. Once ethical progress is understood, the concept allows for a constructive development of ethical truth, one useful in enabling us to characterize inexact core statements we take to be shared among ethically progressive traditions. (249)

Since truth must then be understood in terms of moral progress, we need to look more closely at the latter notion. The pragmatic naturalist does not model moral progress on scientific progress, which he takes to consist in the accumulation of (or, I would say, approximation to) truths. The paradigm is instead technological progress, taken to occur whenever there is a modification leading to an improvement in either function or functional refinement – for example, when a modification

causes an artifact to perform with a lower error rate, or more quickly, cheaply, reliably, etc. (2011: 218). In short, the mark of technological progress is not truth but better performance. Similarly, normative guidance is a social technology with a function that originally was, and still is, the remedying of altruism-failures (2011: 221). Normative guidance, after all, seems to have emerged as a way of responding to the “background problem”: namely, the social instability and conflict caused by altruism-failures confronting early humans. Progress in morality can be measured only in relation to a norm’s performance at fulfilling that original function. It is rare but does take place whenever stable improvements in satisfying that function are introduced. A *progressive* ethical transition occurs in a society whenever there is a change in moral norms resulting in a better performance, over a period of time, in the remedying altruism-failures. A *regressive* transition occurs whenever a change results in worse performance at that function over a period of time. Whenever a norm change results in the same level of performance at the original function of ethics, then neither a progressive, nor a regressive, transition has occurred. Derivatively, we can say some moral beliefs are progressive and others regressive or neutral, depending on their performance in the remedying of altruism-failures over a period of time.

Now the pragmatic naturalist has expounded a conception ethical progress consistent with his historical narrative, one he thinks sufficient to counter the skeptical claim that the evolution of ethics is a history of mere changes. He thereby has the basis for a construal of moral truth in the following non-representational sense: “descriptive counterparts of ethical rules count as true just in case those rules would be adopted in ethical codes as the result of progressive transitions and would be retained through an indefinite sequence of further progressive transitions” (2011: 246). The pragmatic naturalist seems thus in a position to claim that, unlike moral skeptics, he can accommodate moral progress, and, unlike moral realists, he does so without appealing to a notion of moral truth that requires correspondence with moral facts. So it appears that pragmatic naturalism is not vulnerable to the objections that plague either rival.

But pragmatic naturalism still needs to show that, in the end, it is not a quasi-debunking account. For suppose that morality is an evolving social enterprise invented by early humans that was selected by evolution in the way outlined above. A narrative of this sort points to moral truths’ having been neither discovered nor revealed. If so, then neither naturalist (2011: 181 *ff.*) nor non-naturalist forms of realism (2011: 269-71) can be right, since there aren’t, and never were, any external moral truths waiting to be discovered by humans. The anti-realist reasons offered here by the pragmatic naturalist are familiar ones: like the rare episodes of luckily stumbling onto a desired object while sleepwalking, it would be an incredible coincidence that our ancestors somehow stumbled onto certain mind- and language-independent moral truths, so that some of their moral norms turned out to convey true beliefs. Furthermore, given the ethical-project hypothesis, morality is permanently evolving in response to social pressures. So, not only do we continue in the business of inventing moral truths, but (contrary to what some philosophers now

think⁴) there is no hope of ever arriving at a mature folk morality. Moreover, there are no experts who can achieve moral knowledge or determine morality's course of development (2011: 207, 285-86). With the exception of cognitively-impaired people, we are all engaged in the ethical project, though we don't know it. The role of philosophers is merely that of facilitating the conversation about "how we should continue the project of living together" (286).

Given his hypotheses about the evolution and nature of morality, it is quite clear that the pragmatic naturalist thinks moral norms can be completely explain in terms of evolution by natural – and principally cultural (i.e., social) – selection. Yes, he is equipped with a notion of moral truth that, as we saw, is contingent on the more central notion of moral progress. But, unsurprisingly, he cannot invoke the truth of moral beliefs, together with some conception of epistemic moral justification, to claim that we can *know* whether some of our moral norms are true, in the sense of getting things right, independently of their aptitude for the remedying of altruism-failures. All he can say about moral knowledge reduces to this: in light of a plausible evolutionary story, we do know that normative guidance resulted from social pressures. And we do know that some of our moral beliefs are true, in the sense that they perform better than other beliefs in the remedying altruism-failures—though we do *not* know whether they are going to continue to do so, given that ethics is an evolving project (2011: 247-49). But the pragmatic naturalist cannot say that we are justified in thinking that some of moral beliefs are getting things right in the stronger sense needed to amount to moral knowledge if they are also justified. In fact, a discussion of moral knowledge is a glaring omission in this book but entirely consistent with its general skeptical remarks about the impossibility of our ever having it.

To see how pragmatic naturalism does after all amount to an evolutionary quasi-debunking account of morality, consider what it says about moral prescriptions widely accepted across "progressive traditions" such as the norm that one must be honest (i.e., avoid lying, cheating, and the like). Pragmatic naturalism offers an explanation of this norm entirely in terms of its contribution to social cooperation, and thus to the fulfillment of the original function of ethics. Now if that is all there is to explain about 'One must be honest,' the norm can fulfill its evolutionary function, whether it is getting things right (i.e., it is true) or not. Compare the following debunking account of religion, entailed by an evolutionary story holding that religious beliefs have been selected by evolution in ways irrelevant to their getting things right. Our simplified evolutionary story says that theism was evolutionarily selected because of its function in generating positive feelings that help humans cope with existential anxieties created by such factors as the inevitability of death and the constant threat of deception by others.⁵ Those

⁴ For example, Jackson, *From Metaphysics to Ethics*, and Smith, "Moral Realism."

⁵ I owe this way of presenting the analogy with evolutionary debunking of religion to Derek Parfit (NYU seminar, fall 2012). [This short-form note citation to the seminar needs also to appear as a full citation in the bibliography.] My imaginary story about the evolution of religion is broadly inspired by Atran, "Religion's Innate Origins."

who believe that God exists are better able to come to terms with such existential questions, and thus have, on the whole, lives happier than those of non-believers. Suppose that's the background problem religious belief is responding to, and the belief that God exists is performing well at remedying existential anxieties. Further suppose that this is the best explanation of why religious beliefs and practices have developed in all cultures and why people often devote extraordinary time and emotional-energy resources to them. Clearly, if the belief that God exists was selected by evolution for its tendency to help us cope with such existential anxieties, humans would tend to have that belief whether or not the belief is getting things right – i.e., whether or not it is *true* in a stronger sense, which is the one that matters for the evolutionary debunking argument. I submit that the pragmatic naturalist account of 'One ought to be honest' is quasi-debunking in a similar way, for on that account, the adoption of this norm could have amounted to a progressive transition (and be true in the pragmatic naturalist sense), whether or not it is true in the stronger sense of getting things right. In the end, the illusion that pragmatic naturalism can avoid a quasi-debunking conclusion is created by its idiosyncratic notion of truth. But now perhaps that illusion is dispelled. In the upshot, pragmatic naturalism is revealed as only another form of moral skepticism.

NATURALISM IN PHILOSOPHY OF CHEMISTRY OR: WHY DOES METAPHYSICS OF NATURE MATTER?

THOMAS SUKOPP

First of all, it seems rather tricky to relate naturalism to an important field of science which is still philosophically underrated. Although chemistry has long been in the shadow of “guiding sciences” such as physics and biology, it plays a more important role nowadays in the much-fragmented philosophy of science. It is, however, still mostly considered a stepchild of much more respected sciences, such as physics. The aim of this article is to sketch the manifold dimensions of naturalism (Section 1) in connection with the philosophy of chemistry. Second, I will explore how naturalistic accounts in the philosophy of chemistry can be understood by considering that this particular field of science in some cases is special as well as different in comparison with other scientific fields (Section 2). Finally, this article will elaborate a view that tries to reconcile naturalism and culturalism (Section 3).

1. NATURALISM, HISTORY OF SCIENCE, AND PHILOSOPHY OF SCIENCE

At first it may seem puzzling to relate issues of naturalism to the philosophy of chemistry, but some salient questions relating to the latter are closely connected with naturalism (and, of course, to other “isms”). Questions like “Can chemistry be ontologically reduced to physics?” “What is the epistemological status of the chemical concept of elements?” (Paneth, 2003/1962, 113) or “Are chemical laws natural laws in the same notion of ‘law’ like laws in physics?” and others rely more or less on naturalistic or anti-naturalistic assumptions such as scientific realism or the notion of experiments as inventions to discover and extend natural processes.¹

¹ For an overview of topics relating to the philosophy of chemistry, see van Brakel (2014), Scerri (1997, 2013), and Schummer (2006, 2010). For more detailed companions to the philosophy of chemistry, see especially Woody, Findlay, and Needham (Eds.) (2012) and Llored (Ed.) 2013. The neglect or at least underestimation of the philosophy of chemistry in comparison with the philosophy of physics and the philosophy of biology is remarkable and rather strange, since chemists produce much more scientific output in terms of papers than all other scientists. The institutionalization of the philosophy of chemistry is getting better, but it has remained relatively ignored in big conferences on the philosophy of science.

I have to focus on a few topics, especially on the nature of chemical elements and substances, but it might be instructive for the reader to frame an area of research. This broad area encompasses different and sometimes quite distinct disciplines such as the history of science (especially chemistry), the philosophy of science, the sociology of science, the history of philosophy, and various philosophical sub-disciplines. Though it sounds like a truism, a proper culmination of such issues is only to be expected if scholars from these and other disciplines collaborate and rethink traditional concepts. To put it in the words of Bensaude-Vincent, the philosophy of chemistry deserves more attention because it helps to “[...] undermine and more importantly to diversify our metaphysical views of nature and reality” (2014, 60). Again, assuming that there is something to be said in this regard, we can clarify our positions and arguments if we refer to naturalism (see section 1.3). We shall take a closer look at the direct relations between naturalism and prevalent topics in the philosophy of chemistry (Section 1.1), thereby opting for a much closer relation and relevance of the history of science in relation to the philosophy of science (Section 1.2).

1.1 NATURALISM AND ITS RELATIONS WITH PREVALENT TOPICS IN PHILOSOPHY OF CHEMISTRY

According to Putnam (2004, 59), a “[...] very common feature is that, as a rule, ‘naturalism’ is not *defined* [italics in original; TS].” Here, naturalism is—as a working thesis—understood as a philosophical position that asserts a continuous exchange between science and philosophy. This continuity implies that ontological, metaphysical, and methodological questions can partly be answered with references made to science. Maybe—this is one reigning creed of naturalists—there is not a priori philosophy. My own position does not stress that every question relating to the broad field of the philosophy of chemistry should be answered with a naturalistic stance simply because I do not adopt even moderate naturalistic positions regarding relevant topics like “Why is chemistry much more creative than physics?” or “Is chemistry inclined to follow a ethical norms?” (the public image of chemistry carries serious negative connotations relating to pollution and other negative environmental consequences) or “What is the nature of chemical experiment in comparison with experiments in other natural sciences?”

The philosophy of chemistry is one of the recent branches stemming from the philosophy of science, albeit the first book on philosophical issues concerning chemistry is 2,400 years old (see Aristotle for “*De generatione et corruptione*” (in Latin)—“*On generation and corruption*” [English translation]—which means composition and decomposition of compounds, mixtures etc.).² Nevertheless, for a long time, chemistry had to play second fiddle to physics, biology, and other

² See Buchheim (2011) (German translation) or Joachim (no date) for an English translation of this first book that explicitly deals with the philosophy of chemistry. Also, see H.H. Joachim (n. date): *On Generation and Corruption*, Provided by The Internet Classics Archive, see http://classics.mit.edu/Aristotle/gener_corr.html.

so-called “leading” or “guiding” sciences. Naturalism frequently seconds the claim of science X to be a guiding or leading science (see radical versions of naturalism in Section 1.3. Also, see Dawkins or Kutschera for a recent example of the hegemonic attempts of science in a rather ideological way). We might concede that naturalism actually is one reigning creed in the philosophy of physics (see Callebaut 1993, or, more recently, MacArthur and De Caro (Eds.), 2004; Gasser (Ed.), 2007; and Papineau 2015, section 2), but it has not attracted sufficient attention in the philosophy of chemistry. To relate naturalism to the philosophy of chemistry or vice versa does not call for any judgment on naturalism. But it does make sense to argue that the following topics can be dealt with in a naturalistic manner. Questions about homologies and analogous compounds as well as about classificatory division of substances are numerous and cover a wide area. More than this: According to Schummer (2004, 4), s of chemistry should be revised because new, rather exotic species are synthesized—for example, the molecule fragments trapped in argon matrices.

Quite different to physics, chemistry is a classificatory discipline. It creates new compounds, but chemistry does not simply produce new compounds, mixtures, and substances. Its ontology is rich, and encompasses molecules and short-living transitional states, fragments of molecules in mass spectroscopy, or molecule states in argon matrices. The classical nomenclature of chemistry rests on the assumption of a strict correlation between pure substances and molecule. This system needs to be revised because of the above- mentioned rather exotic chemical species. Also, questions about hierarchies of classification are to be reformulated because of such shifts in the meaning of nomenclature.

Epistemological questions: There is no “great unifying theory” in chemistry—for example, as quantum physics in physics. But chemical synthesis—as the chemical and more or less exclusive chemical methods—raises epistemological questions regarding the predictability of the features of new compounds. Since, in contrast to physics, chemistry deals with real substances, and since chemical systems are much more complex than physical systems, chemical models and theories have only medium ranges and scopes.

Further epistemological tasks in the philosophy of chemistry are to analyze the role of experiments and experimentation, and to gain insights into different types of model construction and the corresponding thinking styles (Hacking). In a nutshell, one naturalistic approach is to reconstruct how experiments are *really* conducted and to assume a robust scientific realism (see Section 2.1 for further details).

Methodological questions: Since the philosophy of science in a narrow sense deals with methodological questions like reconstruction, description, modeling and evaluation of scientific methods, and since chemistry is sometimes regarded as the “physics of the outer atomic shell,” reductionist programs have to be analyzed. A reductionist program without doubt makes ambitious naturalistic assumptions (more recently Lombardi and Labarca 2005; van Brakel 2003). One famous locus classicus is Dirac’s dictum: “The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of

chemistry [are] completely known from quantum mechanics” (Dirac, 1929, p.714; cf. Lombardi and Labarca 2005, 126). Another is Rutherford’s harsh view of the priority of physics: “Science, that is either physics or collecting stamps.” It would be rather myopic to simply judge whether reductionism is either definitely successful or a failure (for a brief look at ontological reductionism, see Section 1.3). Quantum chemistry has made enormous progress to explain properties of atoms (see Scerri 1991, 126 and 2004, 94ff.). But quantum physics and quantum chemistry are still far away from being able to explain and to predict the properties and the reaction mechanism of complex chemical systems.

Philosophy of Nature: Should we assume a dichotomy of “natural” vs. “artificial?” The dichotomy of “natural” versus “synthetic” or “artificial” seems to be a descriptive distinction based on ontological considerations. But this is not true. Rather, this separation serves for pseudo-moral reasons or at least for normative reasons to dignify “the natural” and to devalue the “artificial.” Apart from questions of purity, a synthetically created natural flavoring, for example, cannot be distinguished from a flavoring that has been extracted from a plant (Schummer 2003, 728ff.). Now, what does “natural” actually mean? The connection between “natural” and “nature” is tricky, but we should speak about nature (see Section 2.3), and naturalism should not remain silent about nature since it is an “-ism” of nature.

Ethical aspects of philosophy of chemistry and naturalism: In the 19th century the public image of chemistry is fraught with accusations of amorality or anti-morality of chemists. Opinion-makers frequently blamed chemists to act like atheists. It was alleged that chemists preached materialism and were megalomaniacs. Such accusations are overdone, though cum grano salis there is something to be said for this argument: chemists are really creators. They use synthesis (and analysis) as the essential and prevalent methods to produce new compounds and somehow extend the natural range of substances.³

History of chemistry, philosophy of chemistry, and naturalism: Historical demarcations lines of chemistry and physics (not to mention demarcation lines of chemistry and biology) (Nye 1993, 34ff.) shape the formation of chemistry as a scientific discipline, in contrast to the idea of chemistry as an art, as *téchne*, and as an auxiliary science (e.g., for pharmacy and medicine). Historical formation of research paradigms in chemistry evolved within a socio-cultural setting of scientific communities; it influenced laboratory settings and thinking styles (Hacking and Fleck). A naturalistic account aims to reconstruct how scientists in specific historical situations really carried out experiments and how they dealt with ordinary problems faced by chemists such as impure substances and improper instruments.

³ According to Nikolaus Korber (chair of the department of inorganic chemistry, University of Regensburg) each day 15,000 new compounds are synthesized and added to the CAS-register (Chemical Abstracts Service) in Columbus, Ohio, USA.

1.2 NATURALISM, HISTORY OF SCIENCE, AND PHILOSOPHY OF SCIENCE

Another prerequisite for further argumentation is the link between naturalism, the history of science, and the philosophy of science. Independent of this if we favor or refuse naturalism, it has its own history that cannot always be separated from its systematic claims—for example, any reference to materialism points to a long philosophical tradition, beginning with pre-Socratic answers to questions about the ultimate constituents of the universe or about the basic principle underlying nature (Greek: *phýsis*). We shall not confound ancient materialism with modern materialism, but there is a line of argumentation from Democritus (and Leukipp) to the first known philosophers speaking of atoms to materialism in the 18th Century. It is seemingly trivial to revert to many criticisms of an ahistorical view of the philosophy of science, but my argument is that we do not study the history of science just for a better understanding of, for example, Stahl's views on Phlogiston. Systematic and historical questions are intertwined; they are mutually dependent and quite often cannot be properly separated.

Why does the history of science matter for the philosophy of science?

Philip Kitcher's⁴ (2012, 505-524) statement that "[e]pistemology without history is blind" is of some value for our deliberations. Analogously, at least for heuristic purposes, it will be fruitful for understanding the entirely changing views in the philosophy of chemistry to state: "Philosophy of science without history is blind." In a nutshell, Kitcher states a threefold blindness relating to the ahistorical or anti-historical concepts in epistemology: first, the historical roots of philosophical problems are neglected or misconceived; second, to decline the fact that knowledge bearers (scientists and other persons) belong to society (at least to a scientific community) and that knowledge is, at least in part, socially construed or distributed by social concepts; and third, to understand paradigmatic changes in establishment and growth of knowledge, the status quo of knowledge at a given time has to be kept in mind.

The general line of argumentation for a closer systematic linkage of these two disciplines with respect to naturalism follows recent insights into history and philosophy of science-studies (HPS) (Richardson 2014; Stadler 2012; Ash 2012; Giere 1973). But it is also motivated by my conviction of classical philosophical analysis. As Richardson points out, HPS does well in mixing the history of science with the philosophy of science. Of course, the history of science is much more descriptive, while the philosophy of science is partly normative. Since Richardson analyses Carnap's "scientific philosophy," and since Carnap could be called a naturalist (he has a deeply skeptical attitude towards metaphysics), a reconstruction of logical positivism as part of the cultural history of science illuminates its philosophical ambitions. "Second, that understanding the scientific ambitions of logical empiricist philosophy of science can give historians a way of thinking about philosophy of science as something other than a dialectical opponent in the enterprise of understanding

⁴ I am indebted to Fabian Burt for his insightful summary of Kitcher's paper.

science” (Richardson 2008, 89). This concept can quite easily be transferred to other naturalistic programs and elucidates the close relation between the history of science and the philosophy of science. Analyses like Richardson’s (see also Cahn 2002; Ash 2012; Stadler 2012, 231, footnote 63) show that historians and philosophers of science mutually benefit from each other.

This estimation is perhaps too optimistic, but I favor it because it can be supported by many case studies in both the history and the philosophy of chemistry (see Section 2.1).

1.3 NATURALISM: AN OVERVIEW

It is beyond the scope of this paper to elucidate all common naturalistic positions regarding various topics of the philosophy of chemistry. One main line of argumentation in favor of naturalistic assumptions in the philosophy of chemistry is as follows: if we favor a naturalistic ontology, i.e., argue that real atoms and real molecules exist, and then macroscopic characteristics of these molecules can be related to secondary qualities and be understood in terms of naturalistic epistemology. (What do we know about natural things in a framework of naturalistic epistemology? The methodology asks for the significance of natural-scientific methods. This shows, in a nutshell, a relation between ontology, epistemology, and methodology.)

As mentioned above, the dictum of Hilary Putnam can be regarded as a provocation. He states: “A very common feature is that, as a rule, ‘naturalism’ is not defined” (Putnam 2004, 59). We will make an attempt to classify different versions of naturalism. Its claims are to be found on different levels of discussion. I do *not* comment on any kind naturalism here but rather differentiate between the ontological, methodological and epistemological versions of naturalism like many others. Within these levels radical versions are put first, followed by moderate varieties.⁵

ONTOLOGICAL NATURALISM: REDUCTIVE AND NON-REDUCTIVE VERSIONS

Reductive ontological naturalism—as the designation “reductive” denotes—claims that every entity, be it a mental or an immaterial concept or idea, can be reduced to physical or even material objects, things, or corresponding processes. One prominent version is called *Eliminate Materialism*. According to Pauen (1996, 77ff.), it can be summarized by three statements: (1) we have no direct access to our mental states—e.g., the conscious mental state of reading this text; (2) we talk about mental states within folk psychology, which should be replaced by a way of speaking that does not pretend to speak about anything mental; and (3) when folk psychology disappears, we will finally recognize and describe mental states as their true form, namely neural states (this view has

⁵ For more elaborated surveys on naturalism, see Callebaut (1993), Kornblith (1994); Haaparanta (1999). Also, more recently see contribution by Sukopp (2006, 2007), Lewens (2012), and Papineau (2015).

been put forward in early works of Richard Rorty and Paul K. Feyerabend, while the works of Paul and Patricia Churchland also endorse this view).

For our purpose, *physicalism* is understood as the ontological thesis, which implies that our world (the one and only universe) is a physical world, “and that all inhabitants of our world, all things that are existing as real things, are ‘nothing else’ than physical things” (Stich 2000, 105) or everything that exists has a physical nature (in German: “ist physischer Natur”) (Beckermann 2000). Again, we put physicalism into three statements: (1) all things are physical things; (2) all properties are physical properties; and (3) from the first two statements, all events are physical events. Exponents of this position are members of the “Vienna Circle,” such as Rudolf Carnap and Moritz Schlick. In fact, early works of Richard Rorty and Paul K. Feyerabend also support this position.

Materialism is the view that “solely concrete or material things are existing” (Bunge and Mahner 2004, 18), which means that any other entity is at utmost existing fictively. “The world consists solely of things, i. e., concrete or material objects” (Bunge and Mahner, 2004, 21). The exponents of this theory are Mario Bunge, Martin Mahner, and David Armstrong.

Let us briefly discuss non-reductive ontological naturalism. Weak emergentism is compatible with reductionist accounts; it combines physical monism (1), the existence of systemic properties (2) with the thesis of synchronous determinacy (3). This sounds rather complicated and thus I will comment propositions 1 to 3. Here, 1 means that the “substrate” of emergent properties or structure consists solely of material components, while 2 restricts the type of properties that are candidates for emergent properties. Again, 3 says that the properties and dispositions of how the system will act are determined synchronous—i. e., they depend nomologically on its microstructure. Exponents are Mario Bunge, Francisco Varela, and Gerhard Vollmer.⁶

Ontological naturalism matters for the philosophy of chemistry, since most ontological naturalists regard physics as “the” science that tells us what exists. But chemical ontology is much richer and also contributes to problems like reductionism in and of chemistry, the relation between theories of chemical bonds and atomism, and concepts of substance (Scerri 2005).

METHODOLOGICAL NATURALISM

Naturalism is often primarily understood as methodological naturalism: “Naturalistic epistemology is an approach to the theory of knowledge that emphasizes the application of methods, results, and theories from the empirical sciences. It contrasts which approaches that emphasize *a priori* conceptual analysis or insist on a theory of knowledge that is independent of the particular scientific details of how mind-brains work” (Wrenn 2005, 1).

⁶ It is beyond the scope of this contribution to discuss further varieties of ontological naturalism with reference to the mind-body problem, such as anomalous monism, theories of supervenience, strong and weak emergentism, and functionalism (Löffler 1999).

Dirk Koppelberg (2000, 82f.) formulates this species of naturalism as a package of three theses. The first one suggests that philosophy is no adequate basis for science; it rejects any a priori philosophy or “prima philosophia”. The second says that philosophy “has no epistemically privileged point of view compared with science (and humanities)” (Koppelberg 2000). The third formulates and demands the application of scientific investigation and results in philosophy.

Depending on different interpretations of these three statements, we distinguish between three versions of methodological naturalism. The strongest account is called “*eliminative naturalism*” and is closely related with the famous replacement thesis: philosophy has no epistemically privileged point of view. Philosophy has not only to abandon its a priori distinction but should be seen as part of empirical psychology. In this respect, Quine (1969, 293) writes: “Epistemology [...] is only science self-applied.” Further advocates of this theory are Paul and Patricia Churchland.

The *cooperation thesis* is much more moderate: although philosophy has no epistemically privileged point of view, and methodologically starts with everyday concepts of epistemic terms and ideas (Koppelberg 2000, 83), epistemologists “use terms and norms and formulate principles and aims that are not completely included in science” (Koppelberg 2000, 84).

The following *externalistic epistemologies* are quite influential and are sketched out in the following. First, *evolutionary epistemology* supposes that our epistemic ability is part of our evolution, and can be explained by evolutionary processes and principles. We gain knowledge about the world, at least partly, because structures of knowledge have been evolved in adaptation to real structures (of the world). Both efforts and limitations of our cognitive system can be explained in such a way. Exponents of this theory are Donald Campbell and in the German-speaking academic community Rupert Riedl, Gerhard Vollmer, and Franz Wuketits.

Reliabilism (“*substantive naturalism*”) advances the view that a belief is justified (correct or true) if it is generated from a correct working cognitive system, which has “so far has yielded true beliefs, i. e. successful representations of our environment” (Löffler 1999, 56). In other words: a belief is justified by means of its generation. True beliefs can be true beliefs independent of the fact if the believer knows something about the processes actually generating a certain belief (Flonta 2000, 168). The most prominent reliabilist is Alvin I. Goldman.

Finally, *normative naturalism* deals with normative questions without referring to traditional philosophy (of course, normative elements cannot sometimes be neglected in different notions of naturalism. But this is beyond the scope of this essay). The crucial question here is why normative naturalism need not rely only on traditional philosophy. One simple answer is: we can prove empirically whether or not given norms are—in relation to given ends—the appropriate instrument (Hartmann and Lange 2000, 152). This view has been expounded by Larry Laudan, Ronald N. Giere, and Harald I. Brown.

Methodological naturalism is important for understanding its relations with the philosophy of chemistry, since “the” methods of science do just as little exist as it would be a failure to confound the status of experiments in physics with the status of experiments in chemistry in particular, or any other suitable methods of physics and chemistry in general.

STRONG AND WEAK VERSIONS OF NATURALISM: METHODOLOGICAL, EPISTEMOLOGICAL, AND METAPHILOSOPHICAL

I will finally give a brief outline of the programmatic accounts of naturalism situated on the above-mentioned levels.

Radical naturalism vs. Moderate naturalism: According to *radical naturalism*, traditional projects and problems have to be repudiated. In contrast to this, *moderate naturalism* (Tetens 2000, 275) states: “With everything, what human beings are, do and with all their abilities, they are part of the empirically accessible world. This one world is investigated by the single (empirical oriented) sciences.” Exponents of radical naturalism are Willard Van Quine, and Paul and Patricia Churchland. Robert Almeder, Holm Tetens, and Gerhard Vollmer are moderate naturalists.

Expanding Naturalism vs. Integrative Naturalism: Do naturalists work on projects that are not discussed by traditional epistemology? When projects and problems are supplemented [by naturalists], Koppelberg (1996, 75) calls it *expanding* naturalism; when they are modified, he calls it *integrative* naturalism. Expanding naturalists are Hilary Kornblith and Willard Van Quine, while Robert Almeder is an integrative naturalist.

Scientific Naturalism vs. Scientistic Naturalism: How do we proceed to solve problems raised by traditional epistemology? On the one hand philosophers who adopt scientific methods and solutions to solve epistemological problems prefer scientific naturalism. On the other hand philosophers holding the view that epistemological questions can exclusively be solved by science advance the notion of scientistic naturalism (Koppelberg 1996, 76). Most methodological naturalists are scientific naturalists. Again, Willard Van Quine opts for the more radical version—i.e., for scientistic naturalism (“Science itself teaches us”).

Unlimited Naturalism vs. Limited Naturalism: How many traditional projects and questions are still relevant topics to be dealt with? (Grundmann 2001, 9f.) Traditional epistemology advocates the following theses: every person who raises knowledge claims can defend such claims by referring to internal processes such as intuition, memory contents, self-reflection, etc. Furthermore, we have a priori knowledge, and knowledge cannot be reduced to brain processes etc. Unlimited naturalism rejects all theses of traditional epistemology but still carries out research on *all of* its questions. Limited naturalism, however, works on some projects under a naturalistic perspective. Willard Van Quine favors unlimited naturalism, while Robert Almeder is an advocate of limited naturalism.

Within these maps of naturalism topics relating to the philosophy of chemistry can be understood as more or less naturalistically. We will now outline some of these topics.

2. CHEMISTRY: LET’S STUCK TO THINGS? METAPHYSICS, NATURE, AND NATURALISM

This section has three goals: while physics deals with mass points and much more abstract entities than chemistry, it is argued how a naturalistic approach to metaphysics might look like (Section 2.1) by supporting and modifying the

concept of Fritz F. Paneth. Second, we argue why naturalists—assuming that naturalism is not a failure at all—should not abandon metaphysics with respect to the fact that science has to talk about nature (Section 2.2). Finally, what chemistry has to tell us about nature is outlined in Section 2.3.

2.1 WHAT IS NATURALISTIC METAPHYSICS? PANETH ON “BASIC SUBSTANCES” AND “SIMPLE SUBSTANCES”

Science or so-called natural sciences describe, analyze, and predict idealized parts of nature. With respect to chemistry, we favor a notion that thinks of molecules as scientifically accessible parts or components of the real world. We will put forward one concept that has been found in the broad field of the philosophy of chemistry. But first, we have to explain what we understand as “metaphysics.” For this, we first assume that there are at least two different notions of metaphysics: 1) what is existence? Or, what is being? 2) what types of things exist?

To be more precise, Metaphysics is a broad area of philosophy marked out by two types of inquiry. The first aims to be the *most general investigation possible into the nature of reality*: are there principles applying to everything that is real, to all that is? – if we abstract from the particular nature of existing things that which distinguishes them from each other, what can we know about them merely in virtue of the fact that they exist? The second type of inquiry seeks to uncover what *is ultimately real*, frequently offering answers in sharp contrast to our everyday experience of the world. Understood in terms of these two questions, metaphysics is very closely related to ontology, which is usually taken to involve both ‘what is existence (being)?’ and ‘what (fundamentally distinct) types of thing exist?’ [...] (Craig 1998)

For the purposes of this contribution it is sufficient enough to keep in mind that science tries to solve the puzzles that nature gives us and tries to figure out which natural entities, forces, natural laws, among others, exist. One assumption is that the nature of reality can be understood in scientific terms. But even if we define the scope of science vs. the scope of metaphysics or religion, one main point remains: how can nature be separated from a supernatural realm or from any non-scientifically accessible sphere? Furthermore, we can separate different types of metaphysics which, at least from our point of view, indicate the inevitability of metaphysics (Vollmer 2004, 68):

1) the metaphysics of transcendence (primarily occupied with non-empirical terms and non-empirical propositions—see, e.g., Kant’s claim of *synthetic a priori judgments*); 2) the metaphysics of immanence (approximately dealing with ontology, working with empirical and non-empirical terms; Bunge speaks of “exact metaphysics”); and 3) the metaphysics of presuppositions (scrutinizing the operations and results of sciences, and trying to understand its underlying assumptions—i.e., scientific realism); 4) speculative metaphysics (engaged with heuristically valuable speculations).

With respect to chemical entities like molecules, we think that scientific realism is a sound metaphysical view. But, unlike Sankey's conclusion, we do not think that scientific realism is a bundle theory. He thinks that:

[S]cientific realism [...] involve[s] four main tenets: (a) *axiological realism*: the aim of science is truth, and scientific progress consists in advance on that aim; (b) *anti-instrumentalism*: the unobservable entities postulated by scientific theories are conceived as real entities rather than mere predictive devices; (c) *correspondence truth*: truth consists in correspondence between what a statement says about the world and the way the world in fact is; (d) *metaphysical realism*: the world investigated by scientists is an objective reality, the existence and nature of which are independent of human mental activity. (Sankey 2000, 213, footnote vi)

We would like to focus only on the last notion: scientific realism is one kind of metaphysical realism. With respect to the existence of molecules, we argue that '[c]hemistry is a science in which interest is directed towards the secondary qualities of substances' (Paneth 2003/1962, 119). Molecules are manifestations of substances—this is a very old concept that begins with Aristotle. For the moment it is sufficient to think of the Greek word "ousia" as a permanent substratum, "but not with reference to our world of appearance" (Paneth 2002/1963, 122), and rather as formed and determined by chemical laws (think of bonding angles in a molecule or affinities). Laws of nature entirely refer or point to an objective world that is trans-perceptual and trans-empirical. We have to think about this world not in terms of science. Thus, molecules are entities existing in reality; their changes and their recombination to other molecules—through chemical reactions—and their generation and corruption, as mentioned by Aristotle, can be explained in terms of re-arrangement of permanent atoms. According to Paneth, "The secondary qualities exist only in our perceptions, the primary ones pertain to the atoms themselves; in their realm only size, shape, and motion exist" (Paneth 2002/1963, 122).

2.2 WHY NATURALISM IN PHILOSOPHY OF SCIENCE SHOULD NOT ABANDON METAPHYSICS

In short, the answer is: no. To abandon metaphysics would be a metaphysical position, since it would have to be argued why metaphysics should be abandoned. So, it would have to argued metaphysically why metaphysics is useless, superfluous, or simply nonsense. We will outline a revival of substance in relation to elements or the German term "Grundstoff" and try to give reasons as to how to understand the difference between "basic substance" and "simple substance". This parlance shows how scientific realism is somehow akin to naturalistic metaphysics that assumes a real world out there. This real world is, of course, not completely independent of my subjective thinking. "Subjective" here simply means that I am the subject that thinks. This thinking has to refer to something, and this "something" is the structure of the real world. Is there any proof that this world exists? No, but it is highly plausible. An argument for such

plausibility is as follows: most chemists have a stance that they adopt a kind of robust scientific realism.

Now, to come back to our account of a chemical element and the notion of substance: in the words of the eminent chemist-philosopher Paneth:

Nature is assumed to be without qualities, and the properties of colour, sound, taste, smell, etc., are applicable only to our sense representations. The objective prerequisites of the last two qualities are still almost totally obscure; but since even today, as formerly, these qualities play an important part in chemistry, in characterising substances (Paneth 2002/1963, 116).

One naturalistic assumption in the development of the old Greek concept “element” was put forward by Lavoisier (1743–1794), one of the founders of modern chemistry (Bensaude-Vincent 1994). His elements are not “sort of metaphysical causes of the behaviour of bodies, as with the alchemists, but visible and tangible substances with quite definite observable properties” (Paneth 2002/1963 *ibid.*).

The conceptual change of “element” can easily be understood when we compare e.g. Geber’s answer to the question “What is the element sulphur?” to Lavoisier’s answer. Geber would have answered something like: “It is the carrier of mutability by fire.” The advance of Lavoisier’s account of element lies, according to Paneth, in inferring the existence of a basic substance from each simple substance found experimentally. An accompanying disadvantage is that, in the eyes of many, the meaning of “basic substance” was completely hidden by that of “simple substance.”

The contemporary chemist—quite in line with Lavoisier—would answer that sulfur is an element with certain properties such as yellow solid stuff at room temperature, colorless, and without taste and smell. And it would also be mentioned that sulfur cannot be decomposed by chemical methods. The crucial question is:

“[W]hat sense at all is there in saying that the element sulphur is preserved unchanged in its compounds, such as the gaseous, colourless, pungently smelling sulphur dioxide?” (Paneth 2002/1963, 129). The parlance of the element sulfur has a clear meaning when we refer to the element of sulfur as a transcendental principle underlying the phenomena. Hence, the naive-realistic and erroneous assumption is that the properties of sulfur are just those properties which it exhibits to us when it is not combined with other basic substances. We might speak of the properties of a specific substance, as we perceive it. To argue for a modification of Paneth’s view, we quote a longer passage:

But Lavoisier’s definition must not mislead us into applying this simplification to the concept of element, which is used in *two different senses*. I have already attempted to keep apart these different meanings in the above discussion by using the two terms “basic substance” and “simple substance”. I have referred throughout to “*basic substance*” whenever *the indestructible substance present in compounds and simple substances was to be denoted*, and to “*simple substance*” whenever *that form of occurrence was meant in which an isolated basic substance uncombined with any other*

appears to our senses. We cannot ascribe specific properties to an element in the sense of basic substance, since the latter contributes to the generation of the infinitely diverse properties which it exhibits, by itself and in combination with other basic materials; as simple substance it can be characterized by the statement of its properties without prejudice to scientific exactness, as we have earlier shown in general for any substance. In the case of the concept of “simple substance” we may remain naive realists; but in the case of “basic substance”, if we are not to get involved at once in contradictions, we must not overlook the fact that it belongs to a transcendental world devoid of qualities. Even to this day the customary definition follows Lavoisier: “an element is a substance that cannot be decomposed into simpler ones by any chemical procedure”. Thus, to be sure, the assumed *criterion* for an element is that it may be obtained as a simple substance which cannot be decomposed further. But the great *significance* of elements for the whole body of chemical theory lies in the assumption that the substances which produce the phenomenon of “simple substances” serve in the quality-less, objectively real sphere of nature as “basic substances” for the construction not only of simple but also of composite substances. (Paneth 2002/1963, 129f.)

So, a simple substance refers to our phenomenologically and empirically accessible world, whereas the basic substance belongs to a quality-less, objectively real sphere of nature and refers to the substratum, an entity that is persistent in time.

One remaining problem is how to understand the meaning of “substratum” and of “element” in the notion of an eternal entity. Does this lead to Platonism? Not necessarily, since the concept of “element” can be understood without reference to an objective world, where elements exist devoid of *any* properties. Kripke and Putnam have argued “names connect with the world through one single criterion provided by modern chemistry, namely atomic number accompanied by a causal chain from the initial act of baptism” (Scerri 2005, 133).

2.3. HOW CAN “NATURE” BE RELATED TO CHEMISTRY?

Science or natural science deals in many respects with nature, but it often remains silent when we ask how exactly we can grasp a notion of nature—e.g., in physics. The success of science is owed to the fact that many aspects of nature are neglected and can be abandoned simply because of the level of abstraction of modeling and theorizing about nature. Among the big scientific approaches to nature we find a mechanistic, a holistic, and an organizational view of nature.

We restrict ourselves to a brief sketch of the most influential, i.e., the mechanistic, paradigm of nature (Gloy 1996, 99-106). Then, we ask what “nature” actually means in chemistry.⁷

The mechanistic model of nature (abbreviated as MMO) has emerged from early modern science.

⁷ For a more detailed view on metaphysical views of nature in early modern chemistry see Sukopp (2013a, 2013b)

The distinction between artificial vs. natural within the paradigm of MMO is in flux, since a physicist, e.g., can only detect elementary particles as particles that are nonexistent in nature (and thus are somehow artificial). “Nature” is getting a more and more abstract model of selected parts of the world. Owing to complex chemical equipment, chemists can detect radioactive elements, and synthesize pesticides and synthetics. These compounds are rather artificial than natural.

Biologists create “artificial beings”—creatures that do not exist in nature. The existence of compounds, phenomena, creatures’ and other such things is entirely enabled by experimental settings. Thus, “nature” is abstraction-enabled in comparison with the phenomenological-corporal-social⁸ experience of nature. The MMO reduces and restricts “nature” to the parlance of atoms, molecules, and genes as concrete entities. These entities are ontologically and epistemologically reliable as far as they can be described by means of natural laws. The MMO makes the following assumptions:

First, MMO relies on the subject-object dichotomy. Thus, “nature” is “the other,” “the alien,” which has to be subjugated. At least, “Nature” is opposed to the researcher and the experimenter.

Second, “nature” is understood as a (mechanical or biochemical) machine in terms of mechanisms. The machine metaphor allows us to think that “nature” can be described and predicted by mathematics, and that it is ontologically reducible to basic entities.

Third, experiments are designed and carried out to create a non-natural situation in the sense of setting up a certain aspect or feature of nature.

Fourth, “nature” is seen in the dimension of a ruling/serving relation. I cannot elaborate this ethical aspect within this contribution.

Now, what does “nature” mean in chemistry?

First, the term “chemical”, is usually understood as the non-natural and the artificial, especially against the background of an ecological-holistic worldview. The “natural” is more or less value-loaded in a positive sense. The following views on nature from the perspective of pharmaceutical chemistry illustrate a complex web of concepts about nature, which mirror the creative power of chemistry as a discipline that is literally creative (Schummer 2003, 724f.): the molecular design of pharmaceutical substances relies on both experimental synthesis of compounds and computer- simulated design of molecules. In this context, “nature” can be grasped in the following seven dimensions that reflect different notions of nature:

1. Nature is described in the sense that chemistry figures out the possible features of pharmaceutically active substances.

2. Nature is reproduced either a) by extracting substances from natural compounds (plants, animals, or soils) (natural material) or b) compounds are synthesized, i.e., as “artificial” products. “Artificial” here simply means that these compounds are non-natural in the sense that we do not find them in nature.

3. Chemistry learns from nature by isolation of substances (e.g., from extracts of plants).

⁸ The German term “lebensweltlich” would fit better here.

4. Chemistry imitates nature, e.g., by a total synthesis of a substance, which we can also find in extracts of plants.

5. Chemistry improves nature by optimizing a natural compound—i.e., chemists isolate and modify substances.

6. Chemists rival nature by total synthesis of compounds, which does not include chemical reactions of natural extracts, but completely works with non-natural substances.

7. Chemistry controls nature by biosynthesis of genetically-manipulated bacteria.

Some of these views understand that nature can easily be regarded as more or less naturalistically. Reproduction of nature would a kind of ideal. “Nature” in most of the above-mentioned concepts is closely related to a realistic world-view of chemists. Realism, as we have argued, is one of the prevalent naturalistic accounts in the philosophy of chemistry, and it could be argued that everything in nature is exact in the sense in which the chemist deals with nature.

3. CONCLUSION: RECONCILING NATURALISM AND CULTURALISM

So far, we have not really judged naturalism very much. Our conclusion argues that there is a reasonable account that reconciles naturalism and culturalism to understand some of the prevalent topics in the philosophy of chemistry. To avoid common misunderstandings, I do not opt for relativism such as “science as a whole is a construction that can only be judged by one culture” or “science is one of the big narratives,” but I opt for perspectivism that includes a culture-sensitive view. Other disciplines, such the sociology of science, the history of science, or science and technology studies, are more inclined to favor a view that encompasses cultural settings as relevant for theoretical argumentation. Now, why should there be a cultural turn in the philosophy of chemistry? Different academic cultures in different scientific communities located in different countries—within a specific historical setting—shape chemical theories. For example, the vocational and educational practice and the corresponding organizational structures in France and in England from the late 17th century to the climax of the chemical revolution in the 1780s have been analyzed as being crucial for theory formation.

The difference between Lavoisier’s corporative view of knowledge and Priestley’s individualistic epistemology highlights the difference between the institutional organization of French and British science in the late eighteenth century. In the highly organized and centralized community of France, the pressures of formal education, centralized learned societies, employment opportunities, and a competitive system of reward and recognition meant that aspiring French chemist had little choice but to follow the intellectual lead of the academicians in Paris [...] (McEvoy 1988, 210f., cited in. Thagard, “Growth of knowledge”, 5).

Chemistry is much more concerned with substances than other so-called leading sciences such as physics. The metaphysics of scientific realism seems to be the best way to understand what chemistry is all about. The seemingly most scientific science, physics, is much more abstract, and ontologically and methodologically much more distant from nature than chemistry. Scientific realism is akin to helpful metaphysics. Like good metaphysics, scientific realism cannot be proved but can be criticized. Although metaphysics in terms of ontology can be understood as naturalistic, the practice (laboratory work and nomenclature) is more intensely culturally shaped and more than this: theory constructions and the evaluations of correct scientific methods are culturally embedded. Chemistry has always been a laboratory science, and we should keep in mind the word “labor” in “laboratory” to understand that chemistry has evolved as an art and as something much more related to craftsmanship than to science. The thinking style of chemists, to revert to Ian Hacking, is deeply influenced by the socio-cultural setting of laboratories. The problems and objects of the study of chemistry have been provided by and limited by the operations that could be performed on materials in a chemical laboratory. As theoretical structures changed and new objectives supplemented or displaced older ones, the stable setting of the chemical laboratory both identifies chemists and distinguishes them from other natural philosophers who dealt with phenomena (as objects and part of nature) in a much more theoretical and abstract way. The practice of chemistry is as only a physical but also a mental exercise. As chemistry is first and foremost concerned with making individual compounds, it is quite closely connected with culturally influenced parts of science such as laboratory styles and technical equipment. Even if we think that chemistry deals with the transformation of substances, its ontology rests on quite robust assumptions on realism (van Brakel 1997, 253f. und Scerri, 2001).

On the basis of concrete reality, individual substances have led to the view that chemistry is an “impure science.” So, in terms of its relevance to the philosophy of science, chemistry has played second fiddle to physics and other leading or guiding sciences. Paradoxically, its ontology, as outlined above, namely the aspects of the theory of science, leads to no good reputation (see e.g., chemistry as an “impure science” (Bensaude-Vincent 2014, 73, endnote 1) because chemists do not claim to discover the inner secrets of matter. To put it poetically, we may resort to Goethe’s Faust:

Chemistry does not claim to know what the world contains in its innermost heart and finer veins.

But the naturalistic core of chemical ontology is based on the distinction between a “simple substance” built by one chemical element like chlorine (Cl_2). Chemists stick to properties, for example, Chlorine is a green gas with a penetrating smell. But the element chlorine as a basic substance has no perceptible properties like color or smell. The naïve realistic notion can only be applied to chlorine as a simple substance, but as empirically transcendent when we refer to chlorine as a “basic substance.”

In a few words, this is our view of reconciling naturalism with culturalism in an interdisciplinary framework⁹, where we need much more collaboration than so far has been the case.

⁹ One last remark that certainly needs further elaboration points to the relevance of (naturalized) social epistemology. If knowledge bearers are not individuals but groups of scientists or scientific communities interacting with each other and accepting each other more or less as experts, then knowledge certainly is entirely cultural in terms of thinking styles, evaluating knowledge, and accepting or rejecting theories.

BIBLIOGRAPHICAL REFERENCES

ABEL, Günter. *Die Dynamik der Willen zur Macht und die ewige Wiederkehr*. 2. ed. Berlin: de Gruyter, 1998.

ACAMPORA, Christa D. (ed.) *Nietzsche's On the Genealogy of Morals. Critical Essays*. Lanham: Rowman & Littlefield Publishers, 2006.

ALAI, M. *Novel Predictions and the No Miracle Argument*. *Erkenntnis*, 2014, 79 (2): 297-326.

ALGRA, K. *The beginnings of cosmology*. In: LONG, A. A (ed.). *The Cambridge Companion to early greek philosophy*. Cambridge: Cambridge University Press, 1999, p. 45 – 65.

AMERICAN PSYCHIATRIC ASSOCIATION. *Diagnostic and Statistical Manual of Mental Disorders*. Fifth edition. American Psychiatric Association, 2015

ANSCOMBE, G.E.M. *Intention*. Oxford: Oxford University Press, 1957.

ARALDI, Clademir L. *Nilismo, Criação, Aniquilamento. Nietzsche e a filosofia dos extremos*. Col. Sendas & Veredas. São Paulo: Discurso Editorial; Ijuí: Ed. UNIJUÍ, 2004.

ARISTÓTELES. *Metaphysics* (trad. David Ross). In: *The complete works of Aristotle* (The Revised Oxford Translation, J. Barnes, ed.). Vol. 2. Princeton: Princeton University Press, 1984.

ARISTOTLE. *The Works of Aristotle*. Oxford: Oxford University Press, 1921.

ASH, Mitchell G. *Wissenschaftsgeschichte und Wissenschaftsphilosophie – Einführende Bemerkungen*. *Berichte der Wissenschaftsgeschichte*, 2012, 35, 87–98.

ATRAN, S. *Religion's Innate Origins and Evolutionary Background*. In *The Innate Mind*, vol. 2, edited by P. Carruthers, S. Laurence, and S. Stich, 302-317. New York: Oxford University Press, 2006.

AUDI, R. *Moral Knowledge*. Oxford: Oxford University Press, 1997.

AUDI, R. *Philosophical Naturalism at the Turn of the Century*. *Journal of Philosophical Research*, 2000, 25: 27-45.

BAIRD, Davis; SCERRI, Eric. *The Invisibility of Chemistry*. In: BAIRD, Davis; SCERRI, Eric; MCINTYRE, Lee (Eds.): *Philosophy of Chemistry. Synthesis of a New Discipline*. Dordrecht: Springer, 2006, 3-18.

- BAKER, Lynne R. *Naturalism and the First-person Perspective*. New York: Oxford University Press, 2013.
- BANKS, Erik C. *Neutral monism reconsidered*. *Philosophical Psychology*. v. 23, n. 2, April 2010, p. 173–187.
- BARNES, J. *Early greek philosophy*. New York: Penguin Books, 1987.
- BARNES, J. *The presocratic philosophers*. London: Routledge, 1982.
- BECKERMANN, Ansgar. *Ein Argument für den Physikalismus*. In: Keil; Schnädelbach (Eds.), 2000, 128-143.
- BEEBEE, H; SABBARTON–LEARY, N. *Are psychiatric kinds “real”?* *European Journal of Analytic Philosophy*, 2010, 6(1): 112–7.
- BENSAUDE-VINCENT, Bernadette. *Lavoisier: eine wissenschaftliche Revolution*. In: Serres, Michel (Hrsg.): *Elemente einer Geschichte der Wissenschaften*. Frankfurt am Main: Suhrkamp, 1994, 645-685.
- BENSAUDE-VINCENT, Bernadette. *Philosophy of Chemistry or Philosophy with Chemistry?* HYLE - International Journal for Philosophy of Chemistry, 2014, Vol. 20, 59-76.
- BERTINO, Andrea C. *Vernatürlichung*. Monographien und Texte zur Nietzsche-Forschung. Berlin: de Gruyter, 2011.
- BLACKBURN, Simon. *Anti-realist Expressivism and Quasi-realism*. In: COPP, D. *The Oxford Handbook of Ethical Theory*. Oxford: Oxford University Press, 2011, 146–162.
- BLACKBURN, Simon. *Essays in Quasi-Realism*, New York: Oxford University Press, 2006.
- BOGHOSSIAN, Paul. *Fear of Knowledge: Against Relativism and Constructivism*, Oxford: Oxford University Press, 2006.
- BOLTON, D. *What is Mental Disorder? An Essay in Philosophy, Science, and Values*. Oxford: Oxford University Press, 2008.
- BOURGET, David; CHALMERS, David. *What do Philosophers Believe?* *Philosophical Papers*, 2014, 170: 465-500.
- BOYD, R. & Richerson, P.J. *Culture and the evolutionary process*. Chicago: University of Chicago Press, 1985.
- BOYD, R. Homeostasis, species and higher taxa. In R. Wilson, ed., *Species: New Interdisciplinary Essays*. Massachusetts: MIT Press, 1999.

BIBLIOGRAPHICAL REFERENCES

- BOYD, R. *Realism, anti-foundationalism and the enthusiasm for natural kinds*. *Philosophical Studies*, 1991, 61: 127–148.
- BOYER, P. *Intuitive expectations and the detection of mental disorder: A cognitive background to folk-psychiatry*. *Philosophical Psychology*, 2011, 24(1): 95–118.
- BOYER, P. *Religion Explained: The Evolutionary Origins of Religious Thought*. Basic Books, 2001.
- BRADING, K. *A Underdetermination as a Path to Ontic Structural Realism*. In LANRY and RICKLES (eds.). *Structural Realism*. The Western Ontario Series in Philosophy of Science. Volume 77, 2012, pp 99-115.
- BRAKEL, Jaap van. *Philosophy of chemistry between the manifest and the scientific image*. Leuven: Leuven University Press, 2000.
- BRAKEL, Jaap van. *Philosophy of Science and Philosophy of Chemistry*. HYLE – International Journal for Philosophy of Chemistry, 2014, Vol. 20, 11-57.
- BRAKEL, Jaap van. *The Ignis Fatuus of Reduction and Unification*. Back to the Rough Ground. In: *Annals of the New York Academy of Sciences*, 2003, Vol. 988, 30-43.
- BRITO, A. N. *Moral Behavior and Moral Sentiments: On the Natural Basis for Universalism and Egalitarianism*. In: M. Christen et al. (eds.), *Empirically Informed Ethics: Morality between Facts and Norms*. Berlin: Springer, 2014.
- BROWN, T.A.; BARLOW, D.H. *A Proposal for a Dimensional Classification System Based on the Shared Features of the DSM–IV*. *Anxiety and Mood Disorders: Implications for Assessment and Treatment*. *Psychological Assessment*, 2009, 21(3): 256–271.
- BRUNSCHWIG, J.; LLOYD, E. R. *A guide to greek thought*. Harvard: The Belknap Press of Harvard University Press, 2003.
- BRUSOTTI, Marco. *Die Leidenschaft der Erkenntnis. Philosophie und ästhetische Lebensgestaltung bei Nietzsche von Morgenröthe bis Also sprach Zarathustra*. Berlin: de Gruyter, 1998.
- BUCHHEIM, Thomas (Ed. and Transl.). *Aristoteles: Über Werden und Vergehen* (De generatione et corruptione). Hamburg: Meiner, 2011.
- BUCKHOLTZ, J. W.; MAROIS, R. *The Roots of modern justice cognitive and neuronal foundations of social norms and their enforcement*. *Nature Neuroscience*, Vol. 15, N. 5, 2012, May. p. 655, 661.
- BUNGE, Mario; MAHNER, Martin. *Über die Natur der Dinge*. Materialismus und Wissenschaft, Stuttgart: Hirzel, 2004.

- BUNGE, Mario. *Chasing Reality. Strife over reality*. Toronto: University of Toronto Press, 2006.
- BURGE, Tyler. *Mind-Body Causation and Explanatory Practice*. In J. Heil and A. Mele, eds., *Mental Causation*, Oxford: Oxford University Press, 1993, 97-120.
- CAHN, Ralph M. *Historische und Philosophische Aspekte des Periodensystems der Elemente*. HYLE Studies in History and Philosophy of Chemistry, 2002, Vol. 1.
- CALLEBAUT, Werner. *Taking the Naturalistic Turn; or, How Real Philosophy of Science Is Done*. Chicago: University Chicago Press, 1993.
- CAMPIONI, G. *Der französische Nietzsche*. Berlin: de Gruyter, 2009.
- CAMPIONI, G. et al. *Nietzsches persönliche Bibliothek. Supplementa Nietzscheana*. Berlin: de Gruyter, 2003.
- CARNAP, Rudolf. *Der Logische Aufbau der Welt*. Hamburg: Felix Meiner Verlag, 1928.
- CHAKRAVARTTY, Anjan. *A Metaphysics for Scientific Realism. Knowing the Unobservable*, Cambridge: Cambridge University Press, 2007.
- CHALMERS, David. *Facing up to the hard problem of consciousness*. Journal of Consciousness Studies, 1995, 2: 200-219.
- CHALMERS, David. *Phenomenal Concepts and the Explanatory Gap*. In ALTER, T. WALTER, V. (eds.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, Oxford: Oxford University Press, 2005.
- CHALMERS, David. *The Conscious Mind*, Oxford University Press, 1996
- CHISHOLM, R. *Freedom and Action*. In: Freedom and Determinism. Ed Keith Lehrer, 11-44, New York: Random House, 1996.
- CHOWDHURY, A.N. *The definition and classification of Koro*. Culture, Medicine and Psychiatry, 1996, 20(1): 41-65.
- CHURCHLAND, P.M. *The Engine of Reason, the Seat of the Soul*, Cambridge, MA: MIT Press, 1996.
- CHURCHLAND, P.S. *Brain-Wise: Studies in Neurophilosophy*. Cambridge (MA): MIT Press, 2002.
- CHURCHLAND, Patricia S. *Neurophilosophy: Toward A Unified Science of the Mind-Brain*. Bradford Books. Cambridge, MA: MIT Press, 1986.

BIBLIOGRAPHICAL REFERENCES

CHURCHLAND, Paul M. *Matter and Consciousness*. Cambridge, MA: Bradford Books, MIT Press, 1984.

CHURCHLAND, Paul M. *The Engine of Reason, The Seat of the Soul: A Philosophical Journey into the Brain*. Cambridge: MIT Press, 1994.

CIANCIO, Claudio. *Il senso comune nel pensiero ermeneutico*. E. Agazzi (a cura di), *Valore e limiti del senso comune*, Milano: Franco Angeli, 2004, 153-64.

CLARK, L., WATSON, D. REYNOLDS, S. *Diagnosis and classification of psychopathology: Challenges to the current system and future directions*. *Annual Review of Psychology*, 1995, 46: 121–153.

CLARK, Maudemarie. *Nietzsche's Immoralism and the Concept of Morality*. In SCHACHT, Richard. (org.). *Nietzsche, genealogy, morality: essays on Nietzsche's Genealogy of Morals*. Londres: University of California Press, 1994.

CLECKLEY, H. *The Mask of Sanity: An Attempt to Clarify Some Issues About the So-Called Psychopathic Personality*. Mosby, 1941.

COPP, D. Is constructivism an alternative to moral realism? In: BAGNOLI, Carla (ed.) *Constructivism in Ethics*. Cambridge: Cambridge University Press, 2013, 108-132.

COPP, D. *Morality in a Natural World: selected essays in metaethics*. Cambridge: Cambridge University Press, 2007.

COPP, D. *Morality, Normativity, and Society*. New York: Oxford University Press, 1995.

CRAIG, Edward. *Metaphysics*. In: *Routledge Encyclopedia of Philosophy*. Accessed in November 13, 2015.

CRANE, Tim and Mellor, D.H. *There Is No Question of Physicalism*. *Mind*, 1995, 99: 185-206.

CRAVER, C. *Mechanisms and natural kinds*. *Philosophical Psychology*, 2009, 22(5): 575–594.

CUIJPERS, P. (et al). *Preventing the onset of depressive disorders: a meta-analytic review of psychological interventions*. *American Journal of Psychiatry*, 2008, 165(10): 1272–80.

D'IORIO, Paolo. *Le voyage de Nietzsche à Sorrente*. Paris: CNRS Éd., 2012.

DAMASIO, A. *Descartes' Error: Emotion, Reason, and the Human Brain*. Revised Penguin edition, 2005.

- DAMASIO, A. *Looking for Spinoza: Joy, sorrow and the feeling brain*. New York: Harcourt, Inc, 2003.
- DAMASIO, A. *Self Comes to Mind: Constructing the conscious brain*. New York: Pantheon Books, 2010.
- DAMASIO, A. *The feeling of what happens: Body and emotion in the making of consciousness*. New York: Harcourt Brace, 1999.
- DARWALL, S.; GIBBARD, A.; RAILTON, P. *Toward Fin de siècle Ethics: Some Trends*. *Philosophical Review* 101 (1992): 115-189. Reprinted in *Moral Discourse and Practice: Some Philosophical Approaches*. 1997.
- DAVIDSON, D. *Essays on Actions and Events*. 2^a Ed. Oxford: Clarendon Press, 2001.
- DAVIES, M; COLTHEART, M. *Introduction: Pathologies of Belief*. In: COLTHEART, M; DAVIES, M. (eds). *Pathologies of Belief*. New York: Blackwell, 2000.
- DE CARO, M. and VOLTOLINI, A. *Is Liberal Naturalism Possible?* In: De Caro, M; MACARTHUR, D. Eds., 2010: 69-86.
- DE CARO, Mario; MACARTHUR, David. *Introduction: The Nature of Naturalism*. In: De Caro and MACARTHUR, M. 2004: 1-17.
- DE CARO, Mario; MACARTHUR, David. *Naturalism in Question*. Cambridge, MA: Harvard University Press, 2004.
- DE CARO, Mario; MACARTHUR, David. *Normativity and Naturalism*, New York: Columbia University Press, 2010.
- DE CARO, Mario. *Review of J. Ritchie's Understanding Naturalism*. *International Journal of Philosophical Studies*, 19 (4), 527–531.
- DE OLIVEIRA, Nythamar. *Recasting Naturalism and Normativity: Social Constructionism, Reflective Equilibrium and Normative Reconstruction*. In: ARALDI, C., CARMO, J.S., CHAGAS, F.C. (eds). *Naturalism: Contemporary Perspectives*. Pelotas: NEPFil, 2014, p. 163-186.
- DE OLIVEIRA, Nythamar. *Recasting the Naturalism-Normativity Debate: Neuroscience, Neurophilosophy, Neuroethics*. *Principios* v. 20, 2013, p. 79-103.
- DENNETT, D. *Darwin's Dangerous Idea. Evolution and the Meanings of Life*. New York: Simon & Schuster, 1995.
- DENNETT, D. *Freedom Evolves*. New York: Viking, 2003.

BIBLIOGRAPHICAL REFERENCES

- DESCARTES, R. *Discourse on the Method*. Yale University Press, 1996.
- DEVITT, Michael. *Realism and Truth*. Oxford: Blackwell, 1991.
- DEVRIES, Willem A. *Naturalism, the Autonomy of Reason, and Pictures*. *International Journal of Philosophical Studies*, 2010, 18 (3): 395–413.
- DEVRIES, Willem A. *Ontology and the Completeness of Sellars's Two Images*. In: GABBANI, C. (ed.). *Between Two Images. The Manifest and Scientific Conception of Human Being, 50 Years on. Humana Mente*, 2010, 21: 1-18
- DEVRIES, Willem A. *Wilfrid Sellars*. Acumen Publishing and Montreal & Kingston: McGill-Queen's University Press, 2005.
- DIELS, H.; KRANZ, W. *I presocratici* (A cura de G. Reale). Milano: Bompiani, 2014.
- DIÓGENES LAÉRCIO. *Vies et doctrines des philosophes illustres*. Paris: Livre de Poche, 1999 [*Vidas e doutrinas dos filósofos ilustres*. Brasília: UNB, 1988].
- DONINI, P; FERRARI, F. *O exercício da razão no mundo clássico*. São Paulo: Annablume, 2012.
- DUMMETT, Michael. *Truth and Other Enigmas* Cambridge, MA: Harvard University Press, 1978.
- DUMONT, J-P (ed.). *Les présocratiques*. Paris: Gallimard, 1988.
- DUNBAR, Robin I. M. *The Social Brain Hypothesis. Evolutionary Anthropology: Issues, News, and Reviews* Volume 6, Issue 5 (1998): 178–190.
- DUPRÉ, J. *Natural Kinds and Biological Taxa. The Philosophical Review*, 1981, 90(1): 66–90.
- DUPRÉ, J. *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Harvard University Press, 1993.
- DUPRÉ, John. *The Miracle of Monism. In* DE CARO, M., MACARTHUR, D. *Naturalism in Question*. Cambridge, MA: Harvard University Press, 2004.
- ERESHEFSKY, M. *Natural kinds in biology. In*: CRAIG, E. (ed). *Routledge Encyclopedia of Philosophy*. Routledge, 2009.
- ERESHEFSKY, M. *The Poverty of the Linnaean Hierarchy: A Philosophical Study of Biological Taxonomy*. Cambridge University Press, 2001.
- FIELD, Hartry. *Science without Numbers*, Oxford: Blackwell, 1980.

FLANAGAN, Owen; STEPHEN, Martin. *Science and the Modest Image of Epistemology*. *Humana Mente*, 2012, 12: 123-48.

FLONTA, Mircea. *Gemäßigter und radikaler erkenntnistheoretischer Naturalismus*. In: Keil, Schnädelbach (Hg.), 163-186, 2000.

FORREST, Peter. *Why most of us should be scientific realists: A reply to van Fraassen*. *The Monist*, 1994, 77 (1): 47-71.

FORST, R. *Kontexte der Gerechtigkeit, Politische Philosophie jenseits von Liberalismus und Kommunitarismus*. Frankfurt am Main, Suhrkamp, 1994.

FRANKFURT, H. *Freedom of the will and the concept of a person*. *The Journal of Philosophy*. Vo. LXVIII, N. 1, January 14, 1971.

FREEMAN, D., GARETY, P.A. *Delusions*. In FISHER, E., O'DONOHUE, W. (eds). *Practitioners' Guide to Evidence-Based Psychotherapy*. Springer Academic, 2006.

FREGE, Gottlob. *Der Gedanke*. In FREGE, G. *Logische Untersuchungen* (p. 30-53). Göttingen: Vanderhoeck und Ruprecht, 1986.

FRENCH, S. *Identity and Individuality in Classical and Quantum Physics*. *Australasian Journal of Philosophy*, 67: 432-446, 1989.

FRENCH, S. *The structure of the world* Oxford: Oxford University Press, 2014.

FRENCH, Steven. *Structure as a Weapon of the Realist*. *Proceedings of the Aristotelian Society*, 2006, 106: 1-19.

FUSAR-Poli, P., HOWES, O., VALMAGGIA, L., MCGUIRE, P. 'Truman' signs and vulnerability to psychosis. *British Journal of Psychiatry* 193(2): 168, 2008.

GASSER, georg (ed). *How successful is Naturalism?* Heusenstamm bei Frankfurt am Main [u. a.]: Ontos, 2007.

GAZZANIGA, Michael. *The social brain*. New York: Basic Books, 1985.

GIERE, Ronald M. *History and Philosophy of Science: Intimate Relationship or Marriage of Convenience?* In: *British Journal for the Philosophy of Science*, Vol. 24, 282-297, 1973.

GIGON, O. *Los origenes de la filosofia griega*. Madrid: Gredos, 1985.

GILL, M. L.; PELLEGRIN, P. *A Companion to ancient philosophy*. London: Wiley-Blackwell, 2009.

BIBLIOGRAPHICAL REFERENCES

GIORDANO, J., GORDIJN, B. (eds). *Scientific and Philosophical Perspectives in Neuroethics*. Cambridge University Press, 2010.

GIOSAN, C., GLOVSKY, V., HASLAM, N. *The Lay Concept of 'Mental Disorder': A Cross-Cultural Study*. *Transcultural Psychiatry* 38: 317–332, 2001.

GLOVSKY, V., HASLAM, N. *Acculturation and changing concepts of mental disorder: Brazilians in the U.S.A.* *Transcultural Psychiatry* 40: 51–62, 2003.

GLOY, Karen. *Mechanistisches – organizistisches Naturkonzept*. In: Dieselbe (Hg.): *Natur- und Technikbegriffe. Historische und systematische Aspekte: von der Antike bis zur ökologischen Krise, von der Physik bis zur Ästhetik*. Bonn: Bouvier, 1996.

GOETZ, S., TALIAFERRO, C. *Naturalism*. Grand Rapids: William B. Eerdmans Publishing Company, 2008.

GOLD, J., GOLD, I. The “*Truman Show*” *delusion: Psychosis in the global village*. *Cognitive Neuropsychiatry* 17(6): 455–472, 2012.

GOODMAN, N. *Fact, Fiction, and Forecast*. Harvard: Harvard University Press, 1983.

GOODMAN, N. *Ways of Worldmaking*. New York: Hackett, 1978.

GREENE, J., COHEN, J. *For the law, neuroscience changes nothing and everything*. *Phil. Trans. R. Soc. Lond. B* 359,1775–1785, 2004

GRUNDMANN, T. (ed.). *Erkenntnistheorie: Positionen zwischen Tradition und Gegenwart*. Paderborn: Mentis, 2001.

GUTHRIE, W.K.C. *Os filósofos gregos*. Lisboa: Presença, 1987.

HAAPARANTA, Leila. *On the possibilities of naturalistic and of pure epistemology*. In: *Synthese* 118, 31–47, 1999.

HACKING, I. *A tradition of natural kinds*. *Philosophical Studies* 61: 109–126, 1991.

HACKING, I. *Kind Making: The Case of Child Abuse*. In *The Social Construction of What?* Harvard University Press, 1999.

HACKING, I. *Kinds of People: Moving Targets*. *Proceedings of the British Academy* 151: 285–318, 2007.

HACKING, I. *Putnam's theory of natural kinds and their names is not the same as Kripke's*. *Principia* 11(1): 1–24, 2007.

- HACKING, I. *The looping effects of human kinds*. In SPERBER, D., PREMACK, D., PREMACK, J.A. (eds). *Causal cognition: A multidisciplinary debate*. Oxford University Press, 1995.
- HÁJEK, Alan. *Interpretations of Probability*. In ZALTA, E. (ed.) *The Stanford Encyclopedia of Philosophy*. Accessible through the URL: <http://plato.stanford.edu/archives/win2012/entries/probability-interpret/>, 2012.
- HARMAN, G. *Naturalism in Moral Philosophy*. In *Ethical Naturalism: Current Debates*, edited by S. Nuccetelli and G. Seay, 8-23. Cambridge: Cambridge University Press, 2012.
- HARTMANN, Dirk. *Ist der erkenntnistheoretische Naturalismus gescheitert?* In: KEIL; SCHNÄDELBACH (Eds.), 144-162, 2000.
- HASLAM, N. *Categorical vs. dimensional models of mental disorder: The taxometric evidence*. Australian and New Zealand Journal of Psychiatry 37: 696–704, 2003.
- HASLAM, N. *Dimensions of folk psychiatry*. Review of General Psychology 9: 35–47, 2005.
- HASLAM, N. *Kinds of kinds: A conceptual taxonomy of psychiatric categories*. Philosophy, Psychiatry, & Psychology 9(3): 203–217, 2002.
- HASLAM, N. *Natural Kinds in Psychiatry: Conceptually Implausible, Empirically Questionable, and Stigmatizing*. In KINCAID, H., SULLEVIN, J.S. (eds.), *Classifying Psychopathology: Mental Kinds and Natural Kinds*. MIT Press, 2014.
- HASLAM, N., BAN, L., KAUFMANN, L. *Lay conceptions of mental disorder: The folk psychiatry model*. Australian Psychologist 42(2): 129–137, 2007.
- HASLAM, N., GIOSAN, C. *The lay concept of “mental disorder” among American undergraduates*. Journal of Clinical Psychology 58, 479–485, 2002.
- HATFIELD, Gary. *Sense-data and the Philosophy of Mind: Russell, James and Mach*. Principia, v. 6, n. 2, 2002, p. 203-30.
- HAUSER, M. *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong*. New York: Harper Collins, 2006.
- HEIL, John. “Recent Work in Realism and Anti-realism,” *Philosophical Books* 30 (2): 65-73, 1989.
- HERMAN, B. *Moral Literacy*. Harvard University Press, 2007.
- HÖFE, Otfried (ed.). *Zur Genealogie der Moral*. Berlin: Akademie Verlag, 2004.

BIBLIOGRAPHICAL REFERENCES

- HONDERICH, Ted. *How Free Are You? The Determinism Problem*. Oxford: Oxford University Press, 2002.
- HORNSBY, Jennifer. *Simple Mindedness. In Defense of Naïve Naturalism in the Philosophy of Mind*, Cambridge (MA): Harvard University Press, 1997.
- HORWITZ, A.V., WAKEFIELD, J.C. *The Loss of Sadness. How Psychiatry Transformed Normal Sorrow into Depressive Disorder*. Oxford University Press, 2012.
- HULL, D. A matter of individuality. *Philosophy of Science* 45: 335–360, 1978.
- HUME, David. *An Enquiry concerning the Principles of Morals*, edited by Tom L. Beauchamp, Oxford/New York: Oxford University Press, 1998.
- HUSSERL, Edmund. *The Crisis of European Sciences and Transcendental Phenomenology*, Engl. transl. Chicago: Northwestern University Press, 1970.
- HUSSEY. *The beginnings of science and philosophy in archaic Greece*. In: GILL, M. L.; PELLEGRIN, P. *A Companion to ancient philosophy*. London: Wiley-Blackwell, 2009, p. 3 – 19.
- IRWIN, T. *Classical thought*. Oxford: Oxford University Press, 1989.
- JACKSON, A., CAVANAGH, J., SCOTT, J. *A systematic review of manic and depressive prodromes*. *Journal of Affective Disorders* 74(3): 209–17, 2003.
- JACKSON, F. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Clarendon Press, 1998.
- JACOBS, Jon. *Naturalism*. In: Internet Encyclopedia of Philosophy (URL: <http://www.iep.utm.edu/naturali/#H5>), 2009.
- JAEGER, W. *La Teología de los primeros filósofos griegos*. México: FCE, 1992.
- JAMES, W. *The Listening Ebony: Moral Knowledge, Religion and Power among the Uduk of Sudan*. Clarendon Press, 1988.
- JANAWAY, Christopher. *Beyond Selflessness: Reading Nietzsche's Genealogy*. New York: Oxford University Press, 2007.
- JASPERS, K. *General Psychopathology*. Manchester University Press, 1963.
- JOYCE, R. *The Evolution of Morality*. Cambridge, MA: MIT Press, 2006.
- KANNER, L. *Child Psychiatry*. Charles C. Thomas, 1935.
- KANT, I. *Kritik der praktischen Vernunft*. Ed. Weischedel. Darmstadt: WBG, 1998.

KANT, I. *Kritik der reinen Vernunft*. Ed. Weischedel. Darmstadt: WBG, 1998.

KANT, I. *Werke in Sechs Bänden*. Herausgegeben von Wilhelm Weischedel. Wiesbaden: Insel Verlag, 2011.

KEAT, Russell. *Phenomenology and Scientific Realism*. In: HAMMOND, M., HOWARTH, J., KEAT, R. *Understanding Phenomenology*, Oxford: Basil Blackwell, 1991, 271-78.

KEIL, Geert; SCHNÄDELBACH, Herbert (Eds.). *Naturalismus: Philosophische Beiträge*. Frankfurt am Main: Suhrkamp, 2000.

KEKES, J. *The morality of pluralism*. Princeton University Press, 1993.

KELLERT, S.H., LONGINO, H.E., WATERS, K. *Scientific Pluralism*. Minneapolis: University of Minnesota Press, 2006.

KENDLER, K., ZACHAR, P., CRAVER, C. *What kinds of things are psychiatric disorders*. *Psychological Medicine* 41: 1143–1150, 2011.

KERSTING, W. *Liberdade e Liberalismo*. Porto Alegre: Edipucrs, 2005.

KESKINEN, A. *Quine on Objects: Realism or Anti-Realism?* *Theoria* 78 (2): 128-145, 2012.

KIM, Mi Gyung. *Stabilizing Chemical Reality: The Analytic-Synthetic Ideal of Chemical Species*. In: HYLE – International Journal for Philosophy of Chemistry, Vol. 20, 117-139, 2014.

KIRK, G.S., RAVEN, J.E., SCHOFIELD, M. *Os Filósofos Pré-socráticos*. 4.ed. Lisboa: Calouste Gulbenkian, 1994.

KITCHER, P. *Biology and Ethics*. In *The Oxford Handbook of Ethical Theory*, edited by D. Copp, 163-85. Oxford: Oxford University Press, 2006.

KITCHER, Philip. *Epistemology Without History is Blind*. In: *Erkenntnis* 75, 505–524, 2011.

KITCHER, Philip. *The Ethical Project*. Cambridge, MA: Harvard University Press, 2011.

KOPPELBERG, Dirk. *Was ist der Naturalismus in der gegenwärtigen Philosophie?* In: Keil; Schnädelbach (Eds.), 68-91, 2000.

KOPPELBERG, Dirk. *Was macht eine Erkenntnistheorie naturalistisch?* In: *Journal for General Philosophy of Science* 27, 71-90, 1996.

BIBLIOGRAPHICAL REFERENCES

KORNBLITH, Hilary. *Introduction: What is Naturalistic Epistemology?* In: Kornblith, Hilary (Ed.): *Naturalizing epistemology*, Cambridge; London: MIT Press, 1-14, 1994.

KORSGAARD, C. *Realism and Constructivism in Twentieth-Century Moral Philosophy*. In: *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. New York: Oxford University Press, 302-326, 2008.

KORSGAARD, C. *Realismo e construtivismo na Filosofia Moral do século XX*. In: *A Filosofia Prática de Kant: ensaios*. Pelotas: Nephil Online, 2014.

KORSGAARD, C. *The Sources of Normativity*. Cambridge: Cambridge University Press, 1996.

KRIPKE, S. *Naming and necessity*. In D. Davidson & G. Harman, eds., *Semantics of Natural Language*. Reidel, 1972.

KUHN, T.S. *The Structure of Scientific Revolutions*. Chicago University Press, 1962.

LADYMAN, J. *Scientific representation: A long journey from pragmatic to pragmatics*. *Metascience* 20:417-442, 2011.

LADYMAN, J. *What is structural realism?* *Studies in the History and Philosophy of Science* 29, pp. 409-24, 1998.

LADYMAN, J., ROSS, D. *Everything Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press, 2007.

LADYMAN, J., ROSS, D., KINCAID, H. (eds.). *Scientific Metaphysics*. Oxford: Oxford University Press, 2013

LADYMAN, J., ROSS, D., KINCAID, H. (eds.). *Structural Realism*. In: E.N. Zalta, ed., *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/structural-realism/>, 2014.

LAM, D., WONG, G. *Prodromes, coping strategies and psychological interventions in bipolar disorders*. *Clinical Psychology Review* 25(8): 1028–1042, 2005.

LEESER, J., O'DONOHUE, W. *What is a delusion? Epistemological dimensions*. *Journal of Abnormal Psychology* 108: 687–694, 1999.

LEITER, B.; SINHABABU, N. (orgs.). *Nietzsche and Morality*. New York: Oxford University Press, 2007.

LEITER, Brian. *Nietzsche on Morality*. New York: Routledge, 2002.

LENMAN, James. *Moral Naturalism*. *Stanford Encyclopedia of Philosophy*. Link: <http://plato.stanford.edu/entries/naturalism-moral/>, 2006.

LEWENS, Tim. *A Surfeit of Naturalism*. In: *Metaphilosophy*, Vol. 1-2, 46-57, 2012.

LIVESLEY, W. *Diagnostic dilemmas in classifying personality disorder*. In: PHILLIPS, K., FIRST, M., PINCUS, H. (eds). *Advancing DSM: Dilemmas in Psychiatric Diagnosis*. American Psychiatric Association, 2003.

LLORED, Jean-Pierre (Ed.). *The Philosophy of Chemistry: Practices, Methodologies, and Concepts*. Newcastle upon Tyne: Cambridge Scholars Publishing, 2013.

LÖFFLER, Winfried. *Naturalisierungsprogramme und ihre methodologischen Grenzen*. In: QUITTERER, Josef; RUNGGALEDER, Edmund (Eds.): *Der neue Naturalismus – eine Herausforderung an das christliche Menschenbild*, Stuttgart; Berlin; Köln: Kohlhammer, 30-76, 1999.

LOMBARDI, Olimpia; LABARCA, Martín. *The ontological autonomy of the chemical world*. In: *Foundations of Chemistry* 7, 2, 125-148, 2005.

LONG, A. A (ed.). *The Cambridge Companion to early greek philosophy*. Cambridge: Cambridge University Press, 1999.

LOPES, Rogério. *A ambicionada assimilação do materialismo: Nietzsche e o debate naturalista na filosofia alemã da segunda metade do séc. XIX*. In *Cadernos Nietzsche*, n. 29. São Paulo: GEN, 2011, p. 309-352.

LOWE, E.J. *Non-Cartesian Substance Dualism and the Problem of Mental Causation*. *Erkenntnis*, 2006, 65: 5.23.

MACARTHUR, David. *Pragmatism, Metaphysical Quietism, and the Problem of Normativity*. *Philosophical Topics*, 2008, 36 (1): 193-209.

MACDONALD, Graham. *Two Natures: Another Dogma?* In: MACDONALD, G., MACDONALD, C. *John McDowell and His Critics*, Oxford: Blackwell: 2006, 222-234.

MACH, Ernst. *Analyse der Empfindungen*. Jena: Verlag von Gustav Fischer, 1922 [1886].

MAGNUS, B.; HIGGINS, K. (Ed.). *The Cambridge Companion to Nietzsche*. Cambridge: Cambridge University Press, 1996.

MARTIN, R.N.D. *Pierre Duhem. Philosophy and History in the Work of a Believing Physicist*, La Salle: Open Court, 1981

BIBLIOGRAPHICAL REFERENCES

McDOWELL, John. *Meaning, Knowledge and Reality*. London, UK; Cambridge, MA: Harvard University Press, 1998.

MCDOWELL, John. *Mente e Mundo*. Tradução João Vergílio G. Cuter. Ensaio Intr. Hilan Bensusan. Aparecida, SP: Idéias & Letras, 2005 [1994]. (Subjetividade Contemporânea). (Original: McDOWELL, John. *Mind and World*. Cambridge; London: Harvard University Press, 1996 [1994].)

MCDOWELL, John. *Mind and World*, Cambridge (MA): Harvard University Press, 1994.

McDOWELL, John. *Naturalismo na Filosofia da Mente*. Tradução: Sofia I. A. Stein. Veritas, v. 58, n. 3, set./dez. 2013, p. 545-566. (Original: McDOWELL, John. "Naturalism in the Philosophy of Mind" [1999]. In: McDowell, John, *The Engaged Intellect: Philosophical Essays*. Cambridge, MA; London, UK: Harvard University Press, p. 257-275, 2009.)

McDOWELL, John. *Perception as a Capacity for Knowledge*. Millwaukee: Marquette University Press, 2011.

McDOWELL, John. *The Engaged Intellect: Philosophical Essays*. Cambridge, MA; London, UK: Harvard University Press, 2009.

MCDOWELL, John. *Two Sorts of Naturalism*. Reprinted in Id., *Mind, Language, and Reality*, Cambridge, Mass.: Harvard University Press, 1998, 167-97.

MCEVOY, John G. *The Enlightenment and the chemical Revolution*. In: WOOLHOUSE, Roger S. (Ed.): *Metaphysics and philosophy of science in the seventeenth and eighteenth centuries*. Essays for Gerd Buchdahl. Dordrecht: Kluwer, 1988, 307-326.

MCGINN, Colin. *The Mysterious Flame. Conscious Minds in a Material World*, New York: Basics Books, 1999.

MILLIKAN, R.G. *Language Conventions Made Simple*. The Journal of Philosophy XCV no.4., 1998.

MILLIKAN, R.G. *Language, Thought and Other Biological Categories*. Cambridge MA: The MIT Press, 1984.

MOLL, J. *The neural basis of human moral cognition*. Nature Reviews 6 (2005): 799-809, 2005.

MOORE, G. *Nietzsche, Biology and Metaphor*. Cambridge: Cambridge University Press, 2002.

MORAN, Dermot. *Husserl's Crisis of the European Sciences and Transcendental Phenomenology. An Introduction*. Cambridge: Cambridge University Press, 2012, 292-293.

MORGANTI, Matteo. *On the preferability of epistemic structural realism*. *Synthese*, 142: 81–107, 2004.

MÜLLER-LAUTER, Wolfgang. *Nietzsche: seine Philosophie der Gegensätze und die Gegensätze seiner Philosophie*. Berlin: de Gruyter, 1971.

MURPHY, D. *Delusions, Modernist Epistemology and Irrational Belief*. *Mind & Language* 28(1): 113–124, 2013.

MURPHY, D. *Natural Kinds in Folk Psychology and in Psychiatry*. In: H. Kincaid & J. S. Sullevin, eds., *Classifying Psychopathology: Mental Kinds and Natural Kinds*. MIT Press, 2014.

MURPHY, D. *Psychiatry in the Scientific Image*. MIT Press, 2006.

MURPHY, D. *The Folk Epistemology of Delusions*. *Neuroethics* 5(1): 19–22, 2012.

MURPHY, J.W. *Jacques Derrida: A Rhetoric that Deconstructs Common Sense*. *Diogenes*, 1984, 128 (32): 125-140.

MUSGRAVE, Alan. *The ultimate argument for scientific realism*. In: R. Nola (ed.), *Relativism and Realism in Sciences*. Dordrecht: Kluwer Academic Press. 1988, 229–252.

NETA, Ram. *Naturalism in Question*. *Philosophical Review*, 2007, 116: 657-662.

NIETZSCHE, F. W. *Beyond Good and Evil*. Translated by Judith Norman. Cambridge: Cambridge University Press, 2002.

NIETZSCHE, F. W. *Digitale Kritische Gesamtausgabe Werke und Briefe (eKGWB)*. *Digital critical edition of the complete works and letters*, based on the critical text by G. Colli and M. Montinari, Berlin/New York, de Gruyter 1967-, edited by Paolo D'Iorio. <http://www.nietzschesource.org/#eKGWB>.

NIETZSCHE, F. W. *On the Genealogy of Morality*. Translated by Carol Diethe. Cambridge: Cambridge University Press, 2006.

NIETZSCHE, F. W. *Sämtliche Werke. Kritische Studienausgabe (KSA)*. 15 vols. Ed. by Giorgio Colli and Mazzino Montinari. Berlin: de Gruyter, 1988.

NIETZSCHE, F. W. *The Gaya Science*. Translated by Josephine Nauckhoff. Cambridge: Cambridge University Press, 2001.

BIBLIOGRAPHICAL REFERENCES

- NOLAN, Daniel. *Canberra Plan*. In: G. Oppy and N.N. Trakasis eds., *A Companion to Philosophy in Australia and New Zealand*. Victoria: Monash University Publishing, 2010, 98-100.
- NYE, Mary Jo. *From Chemical Philosophy to Theoretical Chemistry. Dynamics of Matter and Dynamics of Disciplines, 1800-1950*. London: University of California Press, 1993.
- O'NEILL, Onora. *Constructivism in Rawls and Kant*. In: FREEMAN, S. (Ed.). *The Cambridge Companion to Rawls*. Cambridge: Cambridge University Press, 2003, 347-367.
- O'SHEA, James. *Wilfrid Sellars: Naturalism with a Normative Turn*. Cambridge: Polity Press, 2007.
- OKASHA, S. *Darwinian Metaphysics: Species and the Question of Essentialism*. *Synthese*, 2002, 131: 191–213.
- OLSON, E. *An Argument for Animalism*. Published in R. Martin and J. Barresi, eds., *Personal Identity*. Blackwell. 2003, p. 318-34.
- PAGE, S. *Mind-independence disambiguated: separating the meat from the straw in the realism/anti-realism debate*. *Ratio*. 2006, 19(3): 321–335.
- PANETH, Fritz A. *The Epistemological Status of the Chemical Concept of Element*. In: *Foundations of Chemistry*. 2003. Vol. 5, 113-147. (Reprinted from the *British Journal for the Philosophy of Science*, Vol. 13, pp. 1–14 (Part I) and 144–160 (Part II), 1962, by permission of Oxford University).
- PAPINEAU, D. *Naturalism*. In: *Stanford Encyclopedia of Philosophy*, 2009, p. 1 – 31 (<http://plato.stanford.edu/archives/spr2009/entries/naturalism/>).
- PAPINEAU, David. *Naturalism*. In: *Stanford Encyclopedia of Philosophy*. <<http://plato.stanford.edu/archives/fall2015/entries/naturalism/>>., 2015.
- PAUEN, Michael. *Mythen des Materialismus. Die Eliminationstheorie und das Problem der psychophysischen Identität*. In: *Deutsche Zeitschrift für Philosophie* 44, 1, S. 1996, 77-100.
- PEREBOOM, Derk. *Free Will, Agency, and the Meaning in Life*. Oxford: Oxford University Press, 2014.
- PORCHER, J.E. *The classification, definition, and ontology of delusion*. *Revista Latinoamericana de Psicopatología Fundamental* 19(1): 2016, 167–181.
- PRICE, Huw. *Naturalism without Representationalism*. In: DE CARO, M. MACARTHUR, D. (eds). 2004, 71-105.

PRINZ, Jesse. *Furnishing the Mind: Concepts and Their Perceptual Basis*. MIT Press, 2002.

PRINZ, Jesse. *Gut Reactions: A Perceptual Theory of Emotion*. Oxford University Press, 2004.

PRINZ, Jesse. *The Conscious Brain*. Oxford University Press, 2012.

PRINZ, Jesse. *The Emotional Construction of Morals*. Oxford University Press, 2004.

PSILLOS, Stathis. *Is structural realism possible?* *Philosophy of Science*, 68 (Supplementary Volume): 2001, S13–S24

PSILLOS, Stathis. *The Scope and Limits of the No Miracles Argument*. In: D. Dieks et al. (eds.), *Explanation, Prediction, and Confirmation*. Series: *The Philosophy of Science in a European Perspective 2*, Dordrecht: Springer, 2011, 23-35.

PSILLOS, Stathis. *The Structure, the Whole Structure and Nothing But the Structure?* *Philosophy of Science*, 2006, 73: 560–570.

PUTNAM, H. *Philosophy in an Age of Science*. In: DE CARO, Mario; MACARTHUR, David (Ed.), Cambridge MA: Harvard University Press, 2012.

PUTNAM, H. *The Collapse of the Fact/Value Dichotomy and Other Essays*, 2002.

PUTNAM, H. *The collapse of the factu/value dichotomy and other essays*. Cambridge: Harvard University Press, 2002.

PUTNAM, H. *The meaning of 'meaning'*. In: *Mind, Language, and Reality*. Cambridge University Press, 1975.

PUTNAM, H. *What Is Mathematical Truth?* In: *Mathematics, Matter and Method. Philosophical Papers*, vol. I, Cambridge: Cambridge University Press, 1975: 60-78.

PUTNAM, Hilary. *The Content and Appeal of "Naturalism"*. In: DE CARO, Mario; MACARTHUR, David (Ed.): *Naturalism in Question*, Harvard: Harvard University Press, 2004, 59-70.

QUINE, W. V. *Replies*. In: DAVIDSON, Donald; HAAPARANTA, Jaakko (Eds.): *Words and Objections*. Dordrecht: Kluwer, 1969, 292-352.

QUINE, W. V. *Reply to Morton White*. In: L. Hahn and P. Schillp (eds.) *The Philosophy of W.V. Quine*, Chicago; Open Court, 1985, 663-665.

QUINE, W. V. *Russell's Ontological Development*. *The Journal of Philosophy*, v. 63, n. 21, American Philosophical Association Eastern Division Sixty-Third Annual Meeting, Nov. 10, 1966, p. 657-667.

BIBLIOGRAPHICAL REFERENCES

- QUINE, W. V. *Two Dogmas of Empiricism*. *The Philosophical Review* 60: 20–43, 1951.
- QUINE, W.V. *Theories and Things*. Cambridge (MA): Harvard University Press, 1981.
- RADDEN, J. *On Delusion*. New York: Routledge, 2011.
- RAWLS, J. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.
- RAWLS, J. *Kantian Constructivism in Moral Philosophy*. In: *Collected Papers*. Ed. S. Freeman. Cambridge: Harvard University Press, 1999, 303-358.
- RÉE, Paul. *Gesammelte Werke (1875-1885)*. Ed. by Hubert Treiber. Berlin: de Gruyter, 2004.
- RICHARDSON, Alan. *Scientific Philosophy as a Topic for History of Science*. In: *Isis*, 2008, Vol. 99, 88-96.
- RUSSELL, B. *On Propositions: What They Are and How They Mean* (1919). In: MARSH, Robert Charles (Ed.). *Logic and Knowledge*. London & New York: Routledge, p. 285-320, 1956.
- RUSSELL, B. *Our Knowledge of the External World*. London: Allen S; Unwin, 1914c.
- RUSSELL, B. *The Analysis of Matter*. London: Kegan Paul, 1927.
- RUSSELL, B. *The Analysis of Mind*. London: George Allen & Unwin Ltd.; New York: The Macmillan Company, 1921.
- RUSSELL, B. *The Nature of Acquaintance* (1914a). In: MARSH, Robert Charles (Ed.). *Logic and Knowledge*. London & New York: Routledge, p. 125-174, 1956.
- RUSSELL, B. *The Philosophy of Logical Atomism* (1918). In: MARSH, Robert Charles (Ed.). *Logic and Knowledge*. London & New York: Routledge, p. 125-174, 1956.
- RUSSELL, B. *The Relation of Sense-data to Physics* (1914b). In: *Mysticism and Logic*. London: Longmans, Green, p. 145-79, 1918.
- RYLE, G. *The Concept of Mind*. Hertford College, University of Oxford, 2009.
- SAATSI, J. *Whence Ontic structural realism? EPSA 2007 Proceedings*, 2009.
- SAMUELS, R. (2009). Delusions as a Natural Kind. In M. Broome & L. Bortolotti, eds., *Psychiatry as Cognitive Neuroscience: Philosophical Perspectives*. Oxford University Press.

SAMUELS, R. & Ferreira, M. (2010). Why Don't Concepts Constitute a Natural Kind? *Behavioral and Brain Sciences* 33(2–3): 222–223.

SANKEY, Howard. *Methodological Pluralism, Normative Naturalism and the Realist Aim of Science*. In: NOLA, Robert; SANKEY, Howard (eds.): *After Kuhn, Popper and Feyerabend. Recent Issues in Scientific Method*. Dordrecht: Springer, 2000, 211-229.

SAVAGE, C. Wade. *Sense-Data in Russell's Theories of Knowledge*. *Minnesota Studies in Philosophy of Science*, v. 12, 1989, p. 138-168.

SCERRI, Eric R. *Chemistry, Spectroscopy, and the Question of Reduction*. In: *Journal of Chemical Education* 68, 2, 1991, 122-126.

SCERRI, Eric R. *Normative and descriptive philosophy of science and the role of chemistry*. In: BAIRD, Davis; SCERRI, Eric; MCINTYRE, Lee (Eds.), 2006, 119-128.

SCERRI, Eric R. *On the Continuity of Reference of the Elements. A Response to Hendry*. In: *Studies in History and Philosophy of Science Part A* 37 (2), 2006b, 309-321.

SCERRI, Eric R. *Philosophy of chemistry: Where has it been and Where is it Going*. In: Llored (Ed.), 2013, 208-225.

SCERRI, Eric R. *Reduction and emergence in Chemistry – Two Recent Approaches*. In: *Philosophy of Science* 74, 2007, 920-931 (originally published in: *Proceedings of Philosophy of Science Association 2006 Meeting*).

SCERRI, Eric R. *Some Aspects of the Metaphysics of Chemistry and the Nature of the Elements*. In: HYLE – *International Journal for the Philosophy of Chemistry* 11, 2, 2005, 127-145.

SCERRI, Eric R.; MCINTYRE, Lee. *The Case for the Philosophy of Chemistry*. In: *Synthese* 111, 1997, 213-232.

SCHACHT, Richard. (ed.). *Nietzsche, genealogy, morality: essays on Nietzsche's Genealogy of Morals*. London: University of California Press, 1994.

SCHACHT, Richard. *Nietzsche*. New York: Routledge, 1983.

SCHACHT, Richard. *Of Morals and Menschen*. In: SCHACHT, Richard. (ed.). *Nietzsche, genealogy, morality: essays on Nietzsche's Genealogy of Morals*. Londres: University of California Press, 1994, 427 - 448.

SCHAFFER, Jonathan. *Monism*. In: E. Zalta, eds., *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/monism/>, 2014.

BIBLIOGRAPHICAL REFERENCES

SCHOPENHAUER, A. *Sobre o fundamento da Moral*. São Paulo: Martins Fontes, 1995.

SCHROER, Robert. *How Far Can the Physical Sciences Reach?* In: *American Philosophical Quarterly*, Vol. 47, No. 3, 2010, 253-266.

SCHUMMER, J.; SPECTOR, T. I.: *The Visual Image of Chemistry: Perspectives from the History of Art and Science*. In: HYLE – International Journal for the Philosophy of Chemistry 13, 2007, 1, 3-41.

SCHUMMER, Joachim. *Philosophie der Chemie: Rück- und Ausblicke*. In: Griesar, Klaus (Hrsg.): *Wenn der Geist die Materie küßt*. Frankfurt: Harry Deutsch, 1-12, http://www.joachimschummer.net/papers/2004pr_PhilChem_Griesar.pdf, 2004.

SCHUMMER, Joachim. *Realismus und Chemie: philosophische Untersuchungen der Wissenschaft von den Stoffen*. Würzburg: Königshausen und Neumann, 1996.

SCHUMMER, Joachim. *The notion of nature in chemistry*. In: *Studies in History and Philosophy of Science*, Vol. 34, 2003, 705-736

SCHUMMER, Joachim. *The Philosophy of chemistry: from Infancy Towards Maturity*. In: BAIRD, Davis; SCERRI, Eric; MACINTYRE, Lee (Eds.), 2006, 19-39.

SCHUMMER, Joachim. *The Philosophy of Chemistry*. In: Allhoff, Fritz (ed.): *Philosophies of Science*. Blackwell-Wiley, 2010, 163-183.

SEARLE, J. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press, 1983.

SEARLE, John. *Can Computers Think?* In: *Minds, Brains and Science*. Reith Lectures. Cambridge, Mass.: Harvard University Press, 1984.

SEARLE, John. *Freedom and Neurobiology. Reflections Free Will, Language, and Political Power*, New York: Columbia University Press, 2007.

SEARLE, John. *The Construction of Social Reality*. New York: Free Press, 1995.

SELLARS, W. *Empirismo e Filosofia da Mente. Com uma introdução de Richard Rorty e um guia de estudos de Robert Brandom*. Tradução de Sofia Inês Albornoz Stein. Petrópolis, RJ: Vozes, 2008. (Coleção Epistemologia) (Tradução de: Sellars, Wilfrid. *Empiricism and the Philosophy of Mind*. Cambridge, London: Harvard University Press, 1997 [1956].)

SELLARS, Wilfrid. *Autobiographical Reflections*. In: H.-N., *Action, Knowledge, and Reality: Critical Studies in Honor of Wilfrid Sellars*, Indianapolis: The Bobbs-Merrill Company, 1975, 277-93.

SELLARS, Wilfrid. *Empiricism and the Philosophy of Mind*; reprinted Cambridge MA: Harvard University Press 1997.

SELLARS, Wilfrid. *Philosophy and the Scientific Image of Man*. In: R. Colodny, ed., *Frontiers of Science and Philosophy*, Pittsburgh: University of Pittsburgh Press, 1962, 35-78.

SEN, A. *A ideia de Justiça*. São Paulo: Companhia das Letras, 2011.

SHAFER-LANDAU, Russ. *Moral Realism: A Defence*, Oxford University Press, 2003.

ŠKODLAR, B., DERNOVŠEK, M.Z., KOČMUR, M. *Psychopathology of Schizophrenia in Ljubljana (Slovenia) From 1881 To 2000: Changes in the Content of Delusions in Schizophrenia Patients Related To Various Sociopolitical, Technical and Scientific Changes*. *International Journal of Social Psychiatry* 54(2): 2008, 101–111.

SMILANSKY, Saul. *Free Will and Illusion*. Oxford: Oxford University Press, 2002.
SMITH, Barry. *The Ontology of Social Reality*. In: B. Smith and John Searle, “An Illuminating Exchange: The Construction of Social Reality,” *American Journal of Economics and Sociology*, 62 (2): 2003, 285-309

SMITH, Barry., CEUSTERS, Werner. *Ontological Realism: A Methodology for Coordinated Evolution of Scientific Ontologies*. *Applied Ontology* 5: 2010, 139–188.

SMITH, M. *Moral Realism*. In: *The Blackwell Guide to Ethical Theory*, edited by H. LaFollette, 15-37. Oxford: Blackwell, 2000.

SPERBER, D. *Explaining culture: A naturalistic approach*. Blackwell, 1996.

SPERBER, D. *Is Symbolic Thought Pre-rational?* In: FOSTER, M., BRANDES, S., eds., *Symbol as Sense*. Academic Press, 1980.

SPERBER, D. *On Anthropological Knowledge*. Cambridge University Press, 1982.

STADLER, Friedrich. *History and Philosophy of Science. Zwischen Deskription und Konstruktion*. In: *Berichte zur Wissenschaftsgeschichte* 35, 2012, 217–238.

STENGERS, Isabelle. *Die doppelsinnige Affinität: Der newtonsche Traum der Chemie im 18. Jahrhundert*. In: Serres, Michel (Hrsg.): *Elemente einer Geschichte der Wissenschaften*. Frankfurt am Main: Suhrkamp, 1994, 527-567.

STICH, Stephen. *Deconstructing the Mind*. New York: Oxford University Press, 1996.

BIBLIOGRAPHICAL REFERENCES

- STICH, Stephen. *Puritanischer Naturalismus*. In: KEIL; SCHNÄDELBACH (Eds.), 2000, 92-112.
- STOMPE, T., FRIEDMAN, A., ORTWEIN, G., STROBL, R., CHAUDHRY, H.R., NAJAM, N., CHAUDHRY, M.R. *Comparison of delusions among schizophrenics in Austria and in Pakistan*. *Psychopathology* 32(5): 1999, 225–234.
- STRAWSON, G. *Freedom and Belief*. Oxford: Clarendon, 1986.
- STRAWSON, Galen. *Freedom and Belief* (revised edition). Oxford: Oxford University Press, 2010.
- STRAWSON, P. *Freedom and Resentment*. In: Proceedings of the British Academy, vol. XLVIII, 1962.
- STREET, S. *A Darwinian Dilemma for Realist Theories of Value*. *Philosophical Studies* 127, 2006: 109-166.
- STREET, S. *Constructivism about Reasons*. In: *Oxford Studies in Metaethics* 3, 2008a, 207– 245.
- STREET, S. *Reply to Copp: Naturalism, Normativity, and the Varieties of Realism worth Worrying About*. In: *Philosophical Issues* 18(1), 2008b, 207–228.
- STREET, S. *The constraints of Hume’s naturalism*. *Synthese*, 152 (3): 2006, 339-51.
- STREET, S. *What is Constructivism in Ethics and Metaethics?* In: *Philosophy Compass* 5, 2010, 363– 384.
- STROUD, Barry. *The Charm of Naturalism*. Reprinted in DE CARO and MACARTHUR, eds. (2004), 1996, 21-35.
- STURGEON, Nicholas L. *Ethical Naturalism*. *The Oxford Handbook of Ethical Theory*. David Copp (Ed.). Oxford: Oxford University Press, 2006, 91-121.
- SUKOPP, Thomas. *How successful is naturalism? Talking about achievements beyond Theism and Scientism*. In: Georg Gasser (ed.): *How successful is Naturalism?* Heusenstamm bei Frankfurt am Main [u. a.]: Ontos, 2007, 77-102.
- SUKOPP, Thomas. *Naturalismus. Kritik und Verteidigung erkenntnistheoretischer Positionen*. Heusenstamm bei Frankfurt u.a.: Ontos, 2006.
- SUKOPP, Thomas. *Robert Boyle, Baconian science, and the Rise of Chemistry in the Seventeenth Century*. In: *Politics and Society* Volume 7, no 2 (special issue), 54-76 (online: <http://www.uvvg.ro/socpol/docs/2013-2/4.%20Thomas%20Sukpp.pdf>), 2013a.

SUKOPP, Thomas. *The Metaphysics of Nature in 17th Century Chemistry, Or Explaining and Explaining Away the Occult*. In: *Conference of the International Journal of Arts & Sciences* Vol. 6 (2), 2013b, 241-248 (online: <http://universitypublications.net/ijas/0602/html/toc.html>).

SWINBURNE, Richard. *Mind, Brain, and Free Will*. Oxford University Press, 2013.

SWINBURNE, Richard. *The Existence of God*. 2 ed. New York: Oxford University Press, 2004.

SZASZ, T. *The Myth of Mental Illness: Foundations of a Theory of Personal Conduct*. Harper & Row, 1961.

TATEYAMA, M., ASAI, M., HASHIMOTO, M., BARTELS, M., KASPER, S. *Transculturalstudy of schizophrenic delusions. Tokyo versus Vienna and Tubingen (Germany)*. *Psychopathology* 31(2): 1998, 59–68.

TETENS, Holm. *Der gemäßigte Naturalismus der Wissenschaften*. In: Keil; Schnädelbach (Eds.), 2000, 273-288.

TUGENDHAT, E. *Lições sobre Ética*. Petrópolis: Vozes, 2010.

TULLY, Robert. *Russell's neutral monism*. *Russell: The Journal of the Bertrand Russell Archives*. v. 8, n. 1, 1988, p. 209-224.

TULLY, Robert. *Three Studies of Russell's Neutral Monism*. *Russell: The Journal of the Bertrand Russell Archives*, v. 13, n. 5, 1993, p. 5-35.

VAN FRAASSEN, Bas. *The Scientific Image*, Oxford: Oxford University Press, 1980.

VAN FRAASSEN. *From a View of Science to a New Empiricism*. In: B. Monton (ed.), *Images of Empiricism: Essays on Science and Stances, with a Reply from Bas C. van Fraassen*, Oxford: Oxford University Press, 2003, 337-83.

VAN Os, J., KAPUR, S. *Schizophrenia*. *The Lancet* 374: 2009, 635–645.

VOLLMER, Gerhard. *Wie viel Metaphysik brauchen wir?* In: WESTERKAMP, Dirk; VON DER LÜHE, Astrid (Eds.): *Metaphysik und Moderne. Ortsbestimmungen philosophischer Gegenwart*. Festschrift für Claus-Artur Scheier. Würzburg: Königshausen & Neumann, 2004, 68-81.

WAKEFIELD, J.C. *Disorder as harmful dysfunction: a conceptual critique of DSM-III-R's definition of mental disorder*. *Psychological Review* 99: 1992, 232–247.

BIBLIOGRAPHICAL REFERENCES

WAKEFIELD, J.C. *The myth of open concepts: Meehl's analysis of construct meaning versus black box essentialism*. *Applied & Preventive Psychology* 11(1): 2004, 77–82.

WATKINS, John. *Scientific Realism versus Common Sense Realism?* In: R.S. Cohen, R. Hilpinen, and Q. Renzong, eds., *Realism and Antirealism in the Philosophy of Science*, Boston Studies in the Philosophy of Science, Dordrecht: Kluwer, 1996, 219-26.

WEISENBERG, M., NEEDHAM, P., HENDRY, R. *Philosophy of Chemistry*. In: ZALTA, E.N. (ed.). *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/chemistry/>. 2011.

WEYL, H. *The Theory of Groups and Quantum Mechanics*. London: Methuen and Co., English trans., 2nd edition, 1931.

WHITE, Stephen L. *Empirical Psychology, Transcendental Phenomenology, and the Self*. In: MARRAFFA, M., DE CARO, M., FERRETTI, F. (eds.). *Cartographies of the Mind. Philosophy and Psychology in Intersection*, Dordrecht: Springer, 2007, 243-54.

WIDIGER, T., LIVESLEY, W., CLARK, L. *An integrative dimensional classification of personality disorder*. *Psychological Assessment* 21(3): 2009, 243–255.

WIDIGER, T., SANDERSON, C. *Toward a dimensional model of personality disorders*. In: LIVESLEY, W. (ed.), *The DSM–IV Personality Disorders*. Guilford, 1995.

WILLARD, Dallas. *The World Well Won. Husserl's Epistemic Realism One Hundred Years Later*. In: D. D. Zahavi and F. Stjernfelt (eds.), *One Hundred Years of Phenomenology*, Dordrecht and Boston: Kluwer Academic Publishers, 2002, 69-78.

WILLIAMS, B. *Moral: Uma introdução à ética*. São Paulo: Martins Fontes, 2005.

WILLIAMSON, Timothy. *The Philosophy of Philosophy*. Blackwell/Brown Lectures in Philosophy. Malden MA: Blackwell Publishing, 2007.

WILSON E. O. *Consilience. The unity of knowledge*. New York: FirstVintage Books Edition, 1999.

WILTSCHE, H.A. *What is Wrong with Husserl's Scientific Anti-Realism?* *Inquiry* 55 (2): 2012, 105-130.

WOODY, Andrea I.; HENDRY, Robin Findlay; NEEDHAM, Paul (Eds.). *Handbook of the Philosophy of Science*. Volume 6. Philosophy of Chemistry. Amsterdam, Boston [u.a.]: Elsevier, 2012.

WORRALL, J. *Structural realism: the best of both worlds?* *Dialectica* 43, 1989, pp. 99-124

WORRALL, John. *Structural Realism: The Best of Both Worlds?* *Dialectica* 43: 1989, 99-124.

WRENN, Chase B. Entry *Naturalistic epistemology*. In: The Internet encyclopedia of Philosophy, <http://www.iep.utm.edu/nat-epis/>, viewed on Nov., 13th 2015.

ZACHAR, P. *A Metaphysics of Psychopathology*. MIT Press, 2014b.

ZACHAR, P. *Beyond natural kinds: Toward a relevant psychiatric taxonomy*. In: KINCAID, H., SULLEVIN, J.S. (eds). *Classifying Psychopathology: Mental Kinds and Natural Kinds*. MIT Press, 2014.

ZACHAR, P. *Psychiatric disorders are not natural kinds*. *Philosophy, Psychiatry, & Psychology* 7(3): 2000, 167–182.

Federal University of Pelotas
Department of Philosophy

DISSERTATIO
FILOSOFIA