Georg Gasser (Ed.)
How Successful is Naturalism?

Publications of the
Austrian Ludwig Wittgenstein Society.
New Series

Volume 4

Georg Gasser (Ed.)

How Successful is Naturalism?

Contents

LIST OF AUTHORS

LYNNE RUDDER BAKER
University of Massachusetts, Amherst (USA), Department of Philosophy.

JOHANNES L. BRANDL
University of Salzburg (AUS), Department of Philosophy.

HELMUT FINK
University of Erlangen-Nürnberg (GER), Department of Theoretical Physics I.

ULRICH FREY
Technical University of Braunschweig (GER), Department of Philosophy.

GASSER GEORG/MATTHIAS STEFAN
University of Innsbruck (AUS), Department of Philosophy (Theological Faculty).

PETER M.S. HACKER
St. John's College, University of Oxford (GB), Faculty of Philosophy.

WINFRIED LÖFFLER
University of Innsbruck (AUS), Department of Philosophy (Theological Faculty).

NANCEY MURPHY
Fuller Theological Seminary (USA), School of Theology.

JOSEF QUITTERER
University of Innsbruck (AUS), Department of Philosophy (Theological Faculty).

MICHAEL C. REA
University of Notre Dame (USA), Department of Philosophy.

THOMAS SUKOPP
Technical University of Braunschweig (GER), Department of Philosophy.

KONRAD TALMONT-KAMINSKI
Marie Curie-Sklodowska University (PL), Faculty of Philosophy and Sociology.

GERD VOLLMER
Technical University of Braunschweig (GER), Department of Philosophy.

Acknowledgments

This volume comprises selected papers of the workshop "How successful is naturalism?" which was organized within the International Wittgenstein Symposium 2006 taking place in Kirchberg am Wechsel, Lower Austria, in the week of August 6th-12th 2006.

Special thanks go to Josef Quitterer, Christian Kanzian, and Joseph Wang who contributed essentially to the success of the workshop.

On behalf of all participants I also thank the staff members at the International Wittgenstein Symposium for their work during the workshop.

I wish to express my gratitude to the Austrian Ludwig Wittgenstein Society (ALWS) and the Austrian Science Funds (FWF) whose generous support made it possible to organize the workshop and to edit this volume.

Introduction

1. THE SUCCESS OF NATURALISM

Without doubt naturalism has been successful in shaping the philosophical landscape like no other philosophical tradition in the second half of the last century. For sure this is the case within analytic philosophy. Jaegwon Kim states it straightforwardly:

> If current philosophy can be said to have a philosophical ideology, it is, unquestionably, naturalism. Philosophical naturalism has guided and constrained analytic philosophy as its reigning creed from much of the twentieth century. (Kim 2003, 84)

David Papineau goes even a step further than Kim by claiming that most contemporary philosophers are not simply naturalists as a matter of fact but

> (…) nearly everybody nowadays *wants* to be a 'naturalist' (…). (Papineau 1993, 1, my italics)

Naturalism is not only the most accepted creed among analytic philosophers but a wide-spread world-view throughout contemporary intellectual culture.

What is so special about naturalism that it became so popular in our times? Does naturalism owe its popularity merely to an intellectual fashion than to any serious argument? Is it chic to be a naturalist? What other reasons than fashion could be there that it took so long for intellectuals to appreciate naturalism?

The relatively recent provenance of naturalism and its influence among philosophical circles in particular and cultural discourse in general can be explained by the rapid development of science. From the dawn of the naturalistic enterprise, a close allegiance with science can be observed. Take, for instance, the volume *Naturalism and the Human Spirit* published by leading American philosophers in 1944. It can be seen as a sort of manifesto for the naturalist movement. A. E. Murphy wrote a detailed review of this volume raising the question about the distinctive philosophical position of naturalists.

He says:

> Starting from the acknowledged achievements of scientific inquiry so far, the 'naturalists' intend to show that these same methods, or others essentially 'continuous' with them, are adequate also to those aspects and dimensions of 'the human spirit' which in the past have often been held on philosophical grounds to transcend the methods and aims of science. (Murphy 1945, 405)

According to Murphy's characterization naturalism appears to be above all a *methodological precept*. It prescribes the scientific method as the only source for obtaining reliable knowledge. Thus, naturalism says that we should understand things by going beyond science as little as possible.

This methodological doctrine rests on an *epistemological thesis* which roughly can be set out as the thesis, that all knowledge we can acquire is obtainable only or foremost through the application of the scientific method. We can put it slightly different as follows: All forms of human investigation are best conducted within the framework of our empirical knowledge of the world and empirical knowledge of our world is paradigmatically gained within scientific discourse.

This epistemological thesis often comes hand in hand with an *ontological thesis* claiming that all that exists is what (in principle) can be studied by science. Science studies the spatio-temporal world. Most naturalists would insist that the whole world is spatio-temporal and all the entities to be found in this world are studied by science. In other words: The entities of our spatio-temporal world are the only inhabitants of reality. This rules out immaterial entities such as God(s), Cartesian souls or spirits. The methodological, epistemological, and ontological ingredients sketched so far are paradigmatic for naturalism. A fourth doctrine can be added which is etiological in nature. In terms of an event causal story it explains how all entities whatsoever have come into existence: Each entity within the spatio-temporal world owes its existence, continuity, and end to the operation of causal forces within the spatio-temporal world. We never go outside the spatio-temporal world for explaining anything which takes place within it. The empirical world which is investigated and explained paradigmatically by the sciences is intelligible in its own right. There is no need to look for any additional explanatory help from outside. Although there is no familiar definition of naturalism at hand, the commitment to an

explanatory closure of the spatio-temporal world can be taken plausibly as a defining feature of naturalism. A consequence of the explanatory closure is that all entities reside within the spatio-temporal world as well. Thus, minimally, naturalism consists in the rejection of supernatural entities and their interactions in our world. Interactions of supernatural entities with our world would break open the explanatory closure of the spatio-temporal world and the ontological commitments coming along with it.

Naturalism was successful in pushing the philosophical discussion into a direction where appeals to theological systems or postulated entities which lay outside the empirical world are considered with great suspicion. The contributors of *Naturalism and the Human Spirit* aimed at approaching those dimensions of the human person in a naturalistic way which so far on philosophical grounds were thought to reside outside the domain of science. These dimensions are, for instance, consciousness, intentionality, and subjectivity. Nowadays most philosophers consider these dimensions of the human person as belonging to the empirical—or to use a more innocent term—to the natural world. As part of the natural world, these phenomena are accessible to scientific investigation. Philosophers embracing explicitly a religious worldview accept the naturalistic intuition that science plays an eminent role in the study of the human person. The Christian philosopher Nancey Murphy, for instance, happily acknowledges that

> all of the human capacities once attributed to the mind or soul are now being fruitfully studied as brain processes—or, more accurately, I should say, processes involving the brain, the rest of the nervous system and other bodily systems, all interacting in a socio-cultural world. (Murphy 2006, 56)

2. NATURALISM AND THEISM

Does this mean that in philosophy the question of naturalism has pretty much been settled—settled in favour of naturalism? It does not. The appearance of a pro-naturalist consensus in contemporary analytic philosophy is quite misleading. It might be true that sciences' achievements are almost unanimously acknowledged. Advocates and critics of naturalism may understand the methods and successful application of science exactly in the same way. They may agree that

philosophy cannot neglect science when it comes to the study of nature and the human person. Naturalists and non-naturalists are, however, not unanimous about the philosophical lessons which should be drawn from sciences' achievements. In turn I will focus on some of those *loci* of philosophical discussion where success and limits of the naturalistic enterprise can most apparently be studied.

Non-naturalists, though sympathetic to naturalist approaches for studying reality, reject the naturalistic commitment of the explanatory and ontological closure of the spatio-temporal world. Science does neither adequately explain nor tell us all what there is and what there is not: There are other realms of reality open for epistemological and ontological reflection which go beyond our empirical knowledge of reality; call it the room of reasons, the realm of the spiritual or the divine. These assumed areas provide further explanatory sources, for instance, for an adequate non-reductive account of understanding (certain forms of) religious experiences within theism. Reductionist tenets of any sort are compatible with the naturalistic program but they are incompatible with theistic views presupposing divine (inter-)action. Naturalistic accounts of religious experience may be valuable by allowing the acceptance and the use of research on the biological, psychological, and social realization of religious experience. However, without an account of divine action, religious experience will be reducible to these merely empirical realms.

There remains a deep disagreement between naturalists and theistic thinkers, although their understanding of science might be compatible. Naturalists treat religious experiences merely as natural phenomena which are adequately describable and explainable with psychological, sociological, or neuro-biological concepts. There is nothing more to add. The exponent of a theistic worldview is an opponent of such a view in the sense that a reduction of religious experiences to mere psychological and socio-cultural explanations is rejected. From a theistic point of view as sketched above for religious experiences a reference to spiritual beings or God is essential.

The particular debate about how religious experiences are adequately conceived points towards a general matter of dispute between naturalism and its theistic rivals. The matter of dispute does not so much concern the achievements of science as such but what conclusions for one's worldview

are drawn from it. Naturalists would say that sciences do not only provide *an* adequate picture of reality but more pretentiously that sciences can provide *the* only adequate picture. Theistic philosophers would disagree. Science does deliver true knowledge of our world but science alone cannot tell us the whole story of our world. More contentiously religious philosophers could even claim that science is of secondary importance for many questions concerning reality, for instance, when it comes to an adequate understanding of the nature of human persons, ethics, or religious experience (Moreland and Rae 2000, 40-47).

To sum it up: The matter of dispute between naturalistic and theistic thinkers concerns the role of science for the understanding of reality *in toto* (Bunge/Mahner 2004, 222-231): Can reality be interpreted in a more coherent and comprehensible way from a naturalistic or theistic point of view? Is naturalism superior to its theistic rival because it has the authority of science behind it? Or does theism dispose of a more comprehensive explanatory power—especially if it takes sciences seriously and aims at harmonizing a scientific understanding of reality with theistic assumptions?

3. NATURALISM AND 'NATURALIZATION'

Though religious worldviews are major rivals of naturalism, the most pressing questions about naturalism do arise from a stance within our empirical world. The great majority of philosophers are secular and have abandoned religion for interpreting reality. Thus, discussion about the evaluation of success and limits of naturalism mostly takes place from a stance within our empirical world.

A major battlefield of contemporary discussion concerns the question how such philosophical key notions as 'self', 'subjectivity', 'first person perspective', 'moral values', 'content', and 'free will' can be accommodated in a naturalistic outlook of reality. Traditionally, these phenomena were understood as residing outside the world accessible to science. Thus, a serious use of these notions requires, according to naturalism, '*naturalization*'. That is, these controversial notions are to be defined from uncontroversial notions of the established sciences. Take, for instance, the problem of the first person perspective: 'Naturalization of the

first person perspective' would mean that our subjective first personal point of view could be reduced to brain activated information processing systems (giving rise to our impression of a first person perspective). The subjective first personal point of view is replaced then with an objective third personal one. There is, however, no agreement about what should count as 'naturalization' and according to which standards it can be achieved.

Furthermore, all the work done by naturalists so far has produced no clear indications how to naturalize the phenomena mentioned above. Francis Crick and Christof Koch state it openly that

> no one has produced any plausible explanation as to how the experience of the redness of red could arise from the actions of the brain. (Crick and Koch 2003, 119)

Although Crick and Koch refer to the problem of qualia only, promising strategies to naturalize subjectivity, content, and the first person perspective in general are missing. Lynne Baker rightly called the first person perspective a test case for naturalism's success (Baker 1998). The lack of successful projects of naturalization, especially in the context of our mental life, nourishes sceptical thoughts about naturalism: Is it not a more promising strategy to treat reasons, meanings, values, or subjectivity as phenomena *sui generis* which can be conceived as powers human beings naturally are endowed with? As phenomena *sui generis* they belong to the 'room of reasons' and ought not to be reduced to something unproblematic from a scientific point of view, such as the realm of causes (McDowell 2004).

McDowell pleas for a liberal form of naturalism in which thinking, knowing and feeling are accepted as being part of our way of being animals. Any aims of naturalizing them by integrating them into the realm of causes and natural laws as the proper space of science are rejected because human beings as rational animals find no place in such a constricted scientific picture anymore. McDowell's plea for liberal naturalism can be interpreted as a sign that restrictive forms of naturalism were unsuccessful so far.

A possible strategy of defence for reductive naturalists might be to provide a modest argument for naturalization. Such an argument would

claim that a sufficient or conclusive reason for satisfactory naturalization cannot be given (yet). What, however, can be given is some reason for thinking that the strategy to naturalize our mental life is the right way to proceed. It is the right way to proceed because all alternatives are less convincing in the light of currently available evidence. All things taken together (current empirical evidence, analysis of arguments…) give a large boost for the assumption that the only practicable strategy consists in naturalizing our mental life and the concepts connected with it (Melnyk 2003, 238-309). Such an argument does not provide a proof for naturalization; it leaves even open whether such a proof can be given. Such an argument aims at a redistribution of the burden of proof. It says that there is no liberty to pretend that naturalization of our mental life is an open question in the sense that non-naturalistic rival theories (for instance a realistic interpretation of our folk psychological concepts or dualistic theories) are equally probable in the light of contemporary scientific knowledge.

Non-naturalists, of course, would disagree with such a conclusion. First, they would claim that the burden of proof lies on the reductionists' side: They have to re-describe and re-explain the phenomena we take for granted in everyday life within a scientific third personal framework. Second, non-naturalists would assume that a change of subject takes place. Phenomena that stimulate our philosophical interests are assimilated with the natural mechanisms which support these phenomena and make them possible.

To sum it up: The concept of 'naturalization' is a major issue of dispute in the current debate on naturalism. It might be worthwhile to study more thoroughly what criteria of success or failure for naturalization are invoked. Such a clarification would be a first step towards setting up standards according to which the status or likelihood of naturalization could be measured.

4. REDUCTIVE AND NON-REDUCTIVE NATURALISM

The discussion whether liberal or more restrictive forms of naturalism shall be pursued points towards another problem discussed in contemporary philosophical debate. The naturalist's paradigm for an adequate explanation of reality is science. Unfortunately it is all but clear what has

to be considered as science. Is a naturalist only committed towards the so-called "hard sciences", physics, and biology, or does he/she have to take into account "soft sciences", such as psychology, as well? This ambiguity gives rise to a whole array of versions of naturalism which are hardly compatible with each other.

There are reductive and non-reductive naturalists. The position of reductive naturalism claims that a complete physics (probably in addition with some other well-established science such as chemistry and biology) would provide all the ontological and explanatory means for understanding reality. Within such a framework, there would not be any need to refer, for instance, to psychological categories for accounting for our mental life. Non-reductive naturalists want to preserve higher level phenomena instead, such as the mental, as a reality *sui generis*, which cannot be reduced to lower levels, for instance, the biological or physical one. They have offered various forms of supervenience relations and developed emergentist scenarios for providing an explanatory and ontological framework within which higher level entities find a natural home in a physical world. According to their understanding restrictive forms of naturalism are too narrow because much of what we consider to be a natural and valuable part of our world is being denied. Furthermore important tools for an adequate understanding of ourselves and our world would be missing, for instance, the causal efficacy of our mental life, responsible agency or a robust understanding of meaning and the normative.

Reductive naturalists tend to accuse their non-reductive colleagues to be rather unclear about how emergentist scenarios or supervenience relations do precisely work: Either these higher level entities are metaphysically inflated and end into some version of dualist thinking; or higher level entities are so tightly bound to their realizing subvenient base that it comes close to a full blown reduction to their physical realizers (Kim 1995). Whether or not a substantial form of physicalism can be combined with the rejection of the so-called higher level reduction is still a great matter of dispute.

In my view, this dispute reflects the need to examine more accurately the epistemological and ontological implications of a layered model or reality as presupposed by most reductive and non-reductive naturalists. As long as the (epistemological and ontological) status of the single layers of

reality and their interrelationship is not untangled and made more precisely deadlocks of the current debates can hardly be solved (Heil 2003, 49-50).

5. NATURALISM, COMMON SENSE, AND VALUES

A distinction similar to reductive and non-reductive naturalists is the one between hard and soft naturalists. P. F. Strawson draws this distinction in *Scepticism and Naturalism*. Hard naturalism, according to Strawson, attempts to view the world in an objective and detached light from the third person perspective. For hard naturalism only natural phenomena as conceived from an impersonal scientific point of view exist. Soft naturalism, on the contrary, expands the notion of existence in such a way that it compromises notions of folk psychology and common sense. Thus, soft naturalism accepts that the world is conceived from a first personal point of view as well. Strawson argues that these two approaches to reality are compatible if they are seen as being valid relative to a certain point of view. In the case of conflict between the two views, Strawson would side with soft against hard naturalism (Strawson 1987, 107). Many naturalists would disagree with Strawson. Drees, for instance, is of the opinion that in the light of the success of science it is rational to give hard naturalism priority over soft naturalism in the case of conflict:

> […] science not only supplements, but, in many instances on good grounds, *corrects*, our (soft) 'natural' understanding of reality. (Drees 1996, 11)

The dispute between hard and soft naturalism points towards the question which epistemological status naturalism assigns to humanities, folk psychology, and common sense. Hard naturalism makes science the prime arbiter of truth: In the case of doubt or dispute science is to be preferred over alternative approaches to reality. Hard naturalism seems to assign full cognitive value, or objectivity, to science alone. Such a campaign arouses fears leading to another form of criticism: If science is the only avenue to a comprehensive theory of the world, then history, poetry, music, and also philosophy have not much to add. If science is in a hegemonic position to pursue objectivity, then the humanities are much farther down or—in the worst case scenario—utterly deficient in achieving objectivity. If such

naturalistic intuitions gain credibility, on the long run, they will have a heavy impact on social life and even on therapeutic application of science itself, for instance in medicine. John Dupré discusses this point shortly on the hand of the treatment of Attention Deficit Disorder Syndrome with the drug Ritalin (Dupré 2004, 53f.). Dupré's argument is not directed against the treatment of problematic cases with psychotropic substances. What he calls attention to is the fact that such a 'scientific and reductive approach' looks natural, if not inevitable from a rational point of view. As a consequence, the success of possible alternative approaches, say a psychological analysis or the study of the child's environment, is evaluated as less promising from the beginning on. It could be argued that the disregard of such complementary or alternative approaches deprives hard naturalism from the very beginning of additional tools of explanation and/or application which might contribute to a therapy's success.

Putnam's critique of hard naturalism seems to be motivated by similar worries but in their thrust they refer to social life in general (Putnam 1990, 142-178). An epistemological demotion of the humanities, art, or common sense will encourage their depreciation. By rejecting any coherent notion of an absolute objective conception of the world Putnam finds himself in a position which is similar to the one of Strawson's soft naturalism: We are beings who cannot have a view of the world that does not reflect our interests and values. Questions about our world are always perspective-dependent. Mathematics, physics, history, art, or poetry show our conceptual choices; the world does not impose one perspective singled out from all others upon us.

What my discussion should have made clear is that hard naturalism is confronted with the reproach of 'de-humanizing' our understanding of reality. If such fundamental concepts of human existence as values, morality, freedom, and subjectivity find no a place in a world conceived from the impersonal view of science, then to be human itself is threatened. Hard naturalists have to find a response to this reproach.

This is not to downgrade the achievements of the naturalistic enterprise. The benefit of naturalism, especially of hard naturalism, was the empirical scrutiny of our common sense view and folk psychology. By asking how well our manifest image can be integrated into the scientific image naturalism draws attention to the defining features, merits, and defects of

our common sense image of the world. Conversely, of course, the role of the scientific image has been subject of close scrutiny as well. To continue to work on these issues is a worthwhile and timely philosophical enterprise since there are no signs that science will be less successful in the years to come. Science will not loose its dominating role in shaping our understanding of reality. Thus, a thorough analysis of the consequences of sciences' achievements for our understanding of reality and its impact on cultural discourse is a significant task for philosophy. The analysis will proceed along the lines I tried to sketch: It will revolve around the explanatory power of alternative worldviews to naturalism, naturalization and reduction, the status of common sense and the humanities in a natural world and the place of values and human interests in a world conceived from the scientific point of view.

This book takes stock of the naturalist debate in recent years. Naturalists and anti-naturalists alike unfold their positions discussing success, failure, and limits of naturalistic approaches. "How successful is naturalism?" makes explicit where the lines of agreement and disagreement between naturalists and their critics are situated in contemporary philosophical discussion. A definite answer regarding naturalism's success and limits will not be found in this book. Clarity about agreement and disagreement between naturalists and non-naturalists alone would be an ample progress however.
We now turn to a summary of the articles.

6. SUMMARY

Gerhard Vollmer, *"Can Everything be Rationally Explained Everywhere in the World?"*: Vollmer belongs to the most prominent German naturalists. The paper contains in a programmatic way the main theses a naturalist has to adopt according to Vollmer's understanding. Guiding principle of his understanding of naturalism is that "everything can be rationally explained everywhere in the world." For putting this principle into practice Vollmer relies on the results of natural science and critical rationalism. The scientific method shall be applied wherever we can apply it. Where we cannot apply it hypotheses must be economical in their ontological postulates and in their explanatory means. Furthermore they are to be

criticizable. The principle of economy and of criticizability tip the balance against the assumption of entities beyond human experience: Souls, angels, or God are imaginable but dispensable for the observation, explanation and interpretation of the world.

Nancey Murphy, *"Naturalism and Theism as Competing Traditions"*: Murphy argues that a typical understanding of the relation between naturalism and theism is that they are two nearly identical worldviews, one with and one without God. Instead, naturalism should be seen as a tradition in its own right, beginning with David Hume and Baron d'Holbach. These intellectuals treat systematically the world as a whole, humanity's place in it, immortality, religion, and the structure of society. Dawkins, Wilson, Dennett, to name a few, are current contributors to this tradition. How is one to compare large-scale traditions of this sort? Murphy draws mainly upon resources from Alasdair MacIntyre to consider what it would take to show this rather new tradition to be rationally superior to its theistic rival. She concludes that naturalism so far was unable to provide a satisfactory account of the moral 'ought' and the foundation of morality. This presents a severe crisis for the naturalist tradition which might be a small step in arguing for the theistic tradition.

Thomas Sukopp, *"How Successful Is Naturalism? Talking about Achievements beyond Theism and Scientism"*: For Sukopp naturalism should not be confounded with "Quine's naturalistic rhetoric." Taking Quine as a paradigm of a naturalistic philosopher amounts to a man of straw in the naturalist/non-naturalist debate. Sukopp holds that naturalism should be understood as being open for scientific and other methods as long as standards of success such as explanatory power, economical use of explanatory means, and capacity of problem solving can be met. Hence a naturalist does not have to adhere dogmatically on natural sciences alone for his resources of explaining and interpreting reality. Because of this tolerance naturalism is not forced to neglect everything beyond physical objects, such as values and norms. Furthermore it is not problematic for naturalism to be modest in its explanatory aspirations: Qualia and free will have not been naturalized yet and maybe they never will. This, however, is no evidence against the success of naturalism. To be successful does not imply to be successful everywhere.

Michael C. Rea, *"How Successful Is Naturalism?"*: According to Rea naturalism suffers from a substantial malady: It commits its advocates to views which are in direct tension with the attitudes, doctrines, and goals which characterize naturalism. According to Rea, naturalists are confronted with a dilemma: If naturalism is characterized as a thesis, then it falls into dissonance because adherence to a thesis is inconsistent with the naturalistic commitment to follow science where it leads. Science might overthrow the thesis which is characteristic for naturalism. This rebuke often can be found among non-naturalists. The second horn of the dilemma is more original and complex: Naturalism is committed to scientific realism and to an ontology including only things accessible to scientific investigation. But the commitment to realism forces naturalists to accept arguments that proceed by way of inference to the best explanation. In doing so, according to Rea, naturalists are forced into an ontology which cannot be investigated by science, namely substance dualism. Then, naturalism is dissonant, if the demand for explanation is rejected, and dissonant if it is accepted.

Ulrich Frey, *"Naturalized Philosophy of Science: A Cognitive Approach"*: Ulrich Frey does not argue for naturalism in general but he gives a concrete example how a progressive naturalistic philosophy might work. His example is a naturalistic philosophy of science based on empirically accessible data about cognitive abilities of scientists. He argues that every investigation of scientific practice needs to consider the cognitive abilities of human beings, including scientists. On the basis of three case studies strong evidence is provided in support of the thesis that sciences like cognitive psychology and evolutionary biology offer good descriptions and explanations of phenomena that are of interest in the philosophy of science. So far most philosophers of science used a coarse grained approach by analyzing scientific paradigms and research programs. Many phenomena, however, are missed that way, because strengths and weaknesses of our individual thinking processes have to be considered as well. Frey's naturalistic approach does not exclude historical and sociological facts. Nor does he aim at naturalizing them. The point he makes is that it is essential for philosophy of science (and for other disciplines in philosophy as well) to rely heavily on natural sciences for methodological and epistemological purposes.

P. M. S. Hacker, *"Passing by the Naturalistic Turn: On Quine's Cul-De-Sac"*: Quine contributed actively to the naturalistic turn away from the a priori methods of traditional philosophy to a conception of philosophy as continuous with natural science. Although there is resistance among naturalists to take Quine as a paradigm of a naturalistic philosopher (see Brandl's and Sukopp's articles), doubtlessly American naturalism is closely associated with Quine. Hacker's contribution is a thorough analysis of Quine's naturalized epistemology. According to Hacker, Quine rarely was concerned with questions of traditional epistemology. However, when he was, his answers were not part of empirically testable theories as he demanded for naturalized epistemology but traditional philosophical claims. Hacker concludes that "naturalized epistemology does not answer the great questions of epistemology and is no substitute for their answers." It remains a major task of epistemological and methodological reflection to point out conceptual confusions and incoherences of scientific theories. This does not imply that philosophy is the Queen of sciences. Rather it should be conceived as a tribunal before which scientific theories may be arraigned when trespassing beyond the limits of their qualification.

Georg Gasser & Matthias Stefan, *"The Heavy Burden of Proof for Ontological Naturalism"*: If one accepts scientific realism and the thesis that explanatory concepts in science imply ontological commitments, then naturalism contains an ontological program as well. The task of a naturalistic ontologist is to draw out the metaphysical implications of contemporary science. Gasser and Stefan attend to this task. First of all, it has to be made clear on which sciences a naturalistic ontologist should rely on. This, however, is all but clear. There is no generally accepted concept of science on the basis of which we can distinguish between acceptable and non-acceptable sciences. The dilemma of naturalism can be subsumed as follows: If almost everything is considered to belong to the scope of science, naturalism becomes so liberal that it runs risk of turning into triviality. If, on the contrary, naturalism becomes more restrictive it leans towards reductive physicalism or eliminativism, a price many philosophers are not willing to pay. An attractive alternative seems to be non-reductive physicalism. Referring to Jaegwon Kim's work Gasser and Stefan argue that this is no viable way either. A consequent form of naturalism seems to

lead towards reductive or eliminative forms of physicalism. Philosophers who are unwilling to bite this bullet do better abandon naturalism.

Konrad Talmont-Kaminski, *"Reason, Red in Tooth and Claw: Naturalising Enlightenment Thinking"*: Talmont-Kaminski's paper starts from the assumption that Enlightenment's conception of rationality became more and more a subject of cynicism in the light of the great historical calamities of the 20th century. Enlightenment's conception of reason was logic-based. Rational solutions were considered to be universal, following necessarily from the information given and they had to conform to appropriate rules, such as logical relations. The main contribution naturalism can make is to bring light into our understanding of what it means for us humans to be rational without falling back to the Enlightenment's extreme on the one hand and the anti-intellectualism of nihilism or fundamentalism on the other. For this task any a priori qualms about reason are to be rejected. A naturalistic account to reason treats epistemic methods as open to development and situated in a specific context. Informed by Peirce's pragmatism Talmont-Kaminski's naturalistic understanding of rationality rejects the view that rational solutions are to be universally valid, rational conclusions have not to follow necessarily from prior information and the primary focus is placed upon actions not upon rules. This modest view of human reason recognizes its frailty and is naturally open for further criticism and development.

Lynne Rudder Baker, *"Naturalism and the First-Person Perspective"*: The first-person perspective poses a challenge to naturalism. Thomas Metzinger has proposed an intriguing account of the first-person perspective that takes up that challenge—an account that draws the consequence that there are no selves, only self-models. Baker uses Metzinger's account as a case study for naturalism. For Baker the first person perspective is essential for the existence of a person. If the first person perspective is irretrievably lost, the person goes out of existence even if the person's body continues to exist. For Metzinger there are no entities in the world that are "selves" or "persons", just self-models. Self-models are products of information-processing systems which are phenomenal in character. We are mistaken to think that our experience of being subjects of experience points towards actual subjects of experience who we are. After a thorough analysis of Metzinger's reductionist account

of the human self Baker works out its semantic, epistemic, and moral consequences. Finally she asks whether it would be rational and even possible to accept such a view as Metzinger exposes it.

Josef Quitterer, *"Which Ontology for Naturalists?"*: In the contemporary discussion of philosophy of mind a major issue is the relationship of folk psychology and scientific explanations of human behaviour. Ontologically folk psychology presupposes the existence of enduring subjects which are the bearers of intentional states. Propositional attitudes presuppose acting and thinking subjects which remain the same during time. Most contemporary naturalists deny that in the world conceived from a scientific point of view there can be *proper physical correlates* for enduring subjects as assumed in folk psychology. The entire folk psychological system and its ontology seem to be incompatible with scientific knowledge. According to Quitterer, however, an analysis of contemporary naturalistic literature in philosophy of mind creates the impression that enduring entities are excluded from the list of possible physical correlates of mental phenomena not so much on scientific grounds but because of a one-sided preference of event ontological accounts. This preference leads to the exclusion of "endurers" from a scientific approach to the human person. Quitterer shows that there are scientific findings about human consciousness and experience which can be interpreted more adequately from the point of view of an ontology of continuants. He concludes that an adequate understanding of the human person needs both—events and continuants. Hence, there are ways to reconcile folk psychological assumptions with current scientific knowledge.

Johannes L. Brandl, *"The Unmysteriousness of Consciousness: A Case Study in Naturalistic Philosophy"*: A naturalistic philosophy of mind is generally associated with physicalist theories. Brandl rejects this link between naturalistic philosophy and physicalistic conclusions drawn out of it. Naturalists are not to be confounded with physicalists. To make this point he uses the problem of consciousness as a case study. Brandl thinks a promising way out of the problematic anti-mentalistic stance many philosophers took after Quine is to return to ontological neutrality as promoted by members of the Vienna Circle. This makes room for what he calls a modest form of naturalism. Such a naturalism is pluralistic from an epistemological and methodological point of view. Everything which can

be explained rationally belongs to the realm of such a modest naturalism. In this sense also consciousness is a natural property: There are reasonable explanations how living creatures come to have conscious experiences. This claim can plausibly be defended against the view that consciousness is mysterious and thus, something non-natural as long as ontology is left out of the game. The problem of qualia in a physical world, mental causation, and the mind-body-problem are simply not part of a modest naturalism's program.

Helmut Fink, „*Indeterminacy of a Free Choice: Ontic, Epistemic, or Logical?*": Fink aims at reconstructing the concept of free will within a naturalistic outlook of reality, that is the doctrine that neural processes like all other processes in nature obey to the laws of physics. He considers three features to be salient for the concept of free will: intelligibility, authorship, and alternative possibilities. Fink argues for a clear distinction between different modes of description on the epistemological level: "Mind talk" is not to be confounded with "matter talk." In addition to modes of description there are also different levels of description: One can either stick to the most elementary building blocks of an entity to be described, or introduce concepts at a more complex level, such as persons. A conceptual reconstruction of "free will" calls for introducing a mental mode and a personal level of description. The bulk of Fink's paper concerns the analysis of various notions of indeterminacy and their importance for the debate of alternative possibilities as presupposition of free will. From a first-person perspective thinking in alternatives is constitutive for a free choice. Fink discusses ontic, epistemic and logical indeterminacy. Fink concludes that it is only epistemic indeterminacy from a first-person perspective, even if restricted to periods of deliberation, which ultimately saves the intuition of alternate possibilities.

Löffler Winfried, *"What Naturalists always Knew about Freedom: A Case Study in Narrative Sources of 'Scientific Facts'"*: In recent years the philosophical landscape in Germany was dominated by one major public debate: the freedom of the will. Well-known brain scientists and empirical psychologists collected ample empirical material that they consider to be strong evidence for determinism. In his article Löffler studies accurately the history of the alleged research and how authors make use of the supposedly abundant empirical material. The conclusion of this study is

rather perplexing: The seemingly robust empirical claims, as they are boasted by dominating scientists of the debate, are flatly wrong. Löffler even speaks of a "piece of neuromythology" which has been created over the years by a mixture of sloppy citations, confidence to hearsay, over-interpretations, slight mistranslations, confabulations, and commingling of probabilistic and strict correlations. Of course, this study is not a refutation of naturalistic accounts of free will. It does not show that we will never be able to solve the problem of free will with the help of empirical investigations. But Löffler's contribution shows that we do good to prove how modern naturalists come to the conclusion that old philosophical puzzles have been solved thanks to modern science.

7. REFERENCES

Baker, R. L. 1998: The First-Person Perspective: A Test for Naturalism. *American Philosophical Quarterly* 35, 327-348.

Bunge, M. and Mahner, M. 2004: *Über die Natur der Dinge*. Stuttgart: Hirzel.

Crick, F. and Koch, C. 2003: A framework for consciousness. *Nature Neuroscience*, 6 (2), 119-126.

Drees, W. B. 1996: *Religion, Science and Naturalism.* Cambridge: Cambridge University Press.

Dupré, J. 2004: "The Miracle of Monism." In: De Caro, M. and Macarthur, D. (eds.) *Naturalism in Question.* Cambridge, MA: Harvard University Press, 36-58.

Heil, J. 2003: *From an Ontological Point of View*. Oxford: Oxford University Press.

Kim, J. 1995: "The Myth of Nonreductive Materialism." In: Moser, Paul K. and Trout J.D. (eds.) *Contemporary Materialism.* New York: Routledge, 133-149.

Kim, J. 2003: The American Origins of Philosophical Naturalism. In: Audi, R. (ed.) *Philosophy in America at the Turn of the Century*. Journal of Philosophical Research, APA Centennial Supplement, 83-98.

McDowell, D. 2004: Naturalism and the Philosophy of Mind. In: De Caro, M. and Macarthur, D. (eds.) *Naturalism in Question.* Cambridge, MA: Harvard University Press, 91-105.

Melnyk, A. 2003: *A Physicalist Manifesto. Thoroughly Modern Materialism.* Cambridge: Cambridge University Press.

Moreland, J. P. and Rae, S. 2000: *Body and Soul. Human Nature & the Crisis in Ethics.* Downers Grove, IL: InterVarsity Press.

Murphy, A. E. 1945: Review of Naturalism and the Human Spirit. *Journal of Philosophy* 42, 400-417.

Murphy, N. 2006: *Bodies and Souls, or Spirited Bodies?* Cambridge: Cambridge University Press.

Papineau, D. 1993: *Philosophical Naturalism.* Oxford: Blackwell.

Putnam, H. 1990: "The Place of Facts in a World of Values." In: Putnam, H. edited by Conant, J.: *Realism with a Human Face.* Cambridge, MA: Harvard University Press, 142-162.

Putnam, H. 1990: "Objectivity and the Science/Ethics Distinction." In: Putnam, H. edited by Conant, J.: *Realism with a Human Face.* Cambridge, MA: Harvard University Press, 163-178.

Strawson, P.F. 1987: *Skeptizismus und Naturalismus.* Frankfurt am Main: Athenäum Verlag.

Part I

Naturalism and Worldviews

Can Everything Be Rationally Explained Everywhere in the World?
Theses and Declarations for Naturalism

Gerhard Vollmer
Technical University of Braunschweig

1. THE BASIS OF DISCUSSION

You can argue about many things: about the meaning of terms, the truth of sentences, the validity of norms and values. But in order to argue at all, you have to agree on some things as well: You need a common language; you have to understand the meaning of assertions, questions, regulations, at least so far as they are important for the discussion; you should know that it is an argument. (Of course not all parties do necessarily *want* this argument; you can also be involved in an argument against your own will.)

Furthermore, for a *meaningful* argument you should agree on what you are arguing about, which means for arguing are allowed and how they can be used. Certain principles of symmetry should be recognized. (If they are complied with, is a different question.) Finally both parties should be able to agree on when an issue is settled.

Certainly you can argue about all that once again: The hierarchy of discussion levels does not have a highest or last level. But the higher you get, the more fundamental the argument gets, and the lesser the points in common.

As scholars we argue foremost about factual issues, secondly about procedural issues as well; but in both cases we assume many things "unquestioned". But as philosophers we examine even and especially such circumstances: We realize them, question them, gather arguments for and against, look for alternatives, define criteria, refer to gaps, circular arguments, contradiction. And certainly we can examine these doings as well. Especially

we inquire about the *conditions* for something being possible, conditions for our talking, asking, concluding, discussing, arguing as well.

The question especially disputed by philosophers is: Which conditions have to be met that perceiving, experiencing, and recognising are possible? In accordance with Kant it is called the *transcendental question*. Kant's question is fruitful, even if his answer is not clear to everybody. But it is extended und applied to other human faculties and activities, especially to scholarly work. What scholars normally assume, in order to be able to work, research and argue, is usually called background assumptions, world view, research leading paradigms, fundamental metaphysical decisions. During scholarly everyday work there is seldom a discussion about them, at best one "philosophizes" about them at the weekend.

In the sense of the division of labour between philosophy and the single disciplines this is absolutely legitimate: Neither on journey nor in the field of research you make progress, if you turn around too often and look back at the starting point. But when you got into a dead end street, it may be useful to return to the starting point, to get your bearings, to change direction, perhaps to choose another starting point. Then many will profit from the fact that philosophers have considered this long ago. In this sense philosophizing is always *thinking in an anticipating manner* as well.

The discussion about naturalism is not an interdisciplinary, but a philosophical one. Admittedly, many scientists are naturalists (in a sense we still have to explain); but most of them are not even familiar with this term, and they would hardly be able to characterize precisely their position or defend it with arguments. Who wants to know if naturalism is tenable, does not go into the laboratory: It is not a question that can be dealt with empirically, even if facts from experience are of—perhaps even crucial—importance.

Now we would like to say as clear as possible what we mean by naturalism. We do this—after a rough characterisation in 2—by formulating and explaining the most important theses of naturalism.

2. TWO IMPORTANT CHARACTERISTICS: UNIVERSALITY AND LIMITATION OF MEANS

We understand naturalism as a natural-philosophical anthropological position. It can be characterized briefly by the thesis: everywhere in the world everything can be explained rationally (*überall in der Welt geht es mit rechten Dingen zu*).[1] According to that, this view distinguishes itself by two characteristics: by its *universal claim* and by the *limitation of means* which are allowed to describe and explain the world.

We are aware that in other contexts „naturalism" can be understood in different senses, e.g. if one talks about the art of naturalism (above all it is about imitation of nature), or when Charles Darwin entitles his work about his voyage around the world with the "Beagle" *A Naturalist's Voyage* (it seems that he only wants to be a *researcher in natural-history*), when Karl Popper talks about naturalistic movements in social sciences (as far as they support the *application of physical methods*), or when in ethics the issue of discussion is naturalism (in a sense in which *norms and values* can be found "*out there*" or can be gained from findable facts—what the naturalism explicated here denies).

The mentioned *claim to universality* of naturalism is essential. Even Kant lets himself be called a naturalist "of a special kind" without protest, perhaps even likes it, for he demands that everything in the sciences should be formulated and explained *naturally*—and not in a *theological* language. (Kant 1788, A 126f.) But at the same time he recognizes a *limit* of science: In order to explain organized beings, especially the *expediency* of organismic structures, *teleological* explanations have to be brought in; a grass root Newton is simply impossible. Only when regarding physics, Kant is a naturalist, but not when regarding biology (and not at all when regarding psychology, epistemology, and ethics). In this case modern naturalism is more ambitious: the unmistakable expediency of organismic structures is explained by the principle of natural selection and therefore ultimately by a causally effective principle. Accordingly, a grass root Newton is possible;

---

[1]   The expression "mit rechten Dingen" is used by Hubert Markl 1983, 75, in order to characterise the attitude of the natural scientist. I first find it as an expression for the naturalistic viewpoint in Winfried Franzen, 1984, 72.

if Charles Darwin has already been the complete Newton or if Gregor Mendel, Ronald Fisher, Julian Huxley, Ernst Mayr, Manfred Eigen and others are needed beforehand, is a question of academic history and—in regard to our problem—of secondary importance. It is crucial that Darwin's theory of selection includes all living beings and therefore biology as a whole within naturalistic explanation, so that teleological explanation becomes dispensable and Thomas Aquinas' or William Paley's teleological proofs of the existence of God lose their persuasiveness. The renunciation of teleology is itself a typical example for the second characteristic of naturalism—the programmatic *limitation of means*. It is not that certain means of description and explanation are prohibited from the outset; it is rather a *principle of economy* according to that the most economical, most simple, and most fundamental hypotheses, theories, models should be preferred among competing, but otherwise equal ones. It is crucial that it is accepted as a principle of selection and as an argument. The before mentioned naturalism's claim to universality certainly goes well with this principle, even if it does not follow compellingly from it.

The expression that everywhere in the world everything can be explained rationally is not very precise. Therefore we want to describe naturalism more exactly, first by presenting its programme. This programme consists of four parts at least[2]:

(1)   It demands and draws a *cosmic overall picture*, a "conception of the world".

(2)   It also ascribes a certain place in the universe to *man*.

(3)   It includes in its claims and approaches for explanation *all* human faculties, language, recognition, academic research, moral action, and aesthetic judgment as well.

(4)   On this basis it demands and develops especially a naturalistic anthropology,

---

[2]   For the naturalistic programme see Ernest Nagel, "Naturalism Reconsidered", *Proceedings and Addresses of the American Philosophical Association*, *28*, October 1955, 5-17. Nagel does not speak of a programme, but of an "extensive intellectual image of nature and man"; it includes "a general design for the cosmic events and for man's position within and logic of research".

a naturalistic epistemology,
a naturalistic methodology of research,
a naturalistic ethics,
a naturalistic aesthetics.

Within the context of this programme naturalism holds theses of a sort that regards content and methodology.

3. THESES OF NATURALISM

(i)   Only as much metaphysics as necessary!

The opinions about metaphysics differ very much. The traditional philosophy was rather pro-metaphysics. In contrast to that, positivism, instrumentalism, pragmatism, logical empiricism, and the Vienna Circle were extremely anti-metaphysics. But it became obvious that we cannot do without metaphysical assumptions, not even in scholarly studies and science. Therefore naturalism does not reject metaphysics completely, does not think of it as inferior, but tries to differentiate it from empirical science, e.g. by Popper's demand for falsifiability. In any case, differentiating does not mean abolishing, as Popper is sometimes accused of doing.

Then, how much metaphysics should we permit? The naturalistic answer is unambiguous: only as much metaphysics as necessary—necessary for research, for progress of knowledge, for life. Accordingly, the naturalist looks for a sort of minimum-metaphysics (see Wendel 1993, 104). It includes the assumption of a world being independent of consciousness, structured, and related (see iii, viii, ix, xi) and its partial recognisability by perception, experience and an intersubjective science (see iii, viii, ix, xi). This view is also called "hypothetical realism".

Although such metaphysical premises cannot be checked empirically, they are nevertheless open to criticism, e.g. in regard to freedom of contradiction, value of explanation, self-applicability, freedom of arbitrariness, intellectual economy, prolificness. And if they are open to criticism, there can be good reasons for rescinding them and replacing them by others. The naturalist's minimum-metaphysics is put—occasionally— under the microscope of rational critique.

Doubt is here—as with Descartes—deliberation, not an existential mood. Nothing is undeceivable not our own viewpoint, not our minimum metaphysical premises, not even language with which we formulate everything, not even our own questions, our own doubts. But that does not mean that the naturalist leaves everything open. Certainly, the naturalist has convictions, certainties, convincingness; but he is also conscious of their fallibility.

In regarding such a minimum-metaphysics two points can be differentiated: How can it be found? And of what importance is it? It is found by analysis and reflection upon the premises of our thinking and acting. This analysis is a typical philosophical activity. It examines our linguistic usage, performances and mistakes of our perception and our experience, but scientific methods and results as well. The role of our minimum-metaphysics is to guide our thinking and acting. We simply cannot do completely without such epistemological and action guiding assumptions and maxims. The extent of our minimum-metaphysics follows from what we want to know or what we want to do.

(ii)   As much realism as possible!

No one can be forced into realism by arguments. Even the most radical solipsist who thinks only his momentary consciousness to be existent is irrefutable. Although his position is not plausible, but it is—when formulated prudently—non-circular and non-contradictory, logically consistent and modest. That is why Schopenhauer compares aptly the solipsist with a maniac in an impregnable log cabin.

But there are good reasons in favour of realism (Vollmer 1993a, 161-181). Certainly they are not of a logical, nor an empirical, nor an historical nature, but of a *meta-theoretical* nature. In contrast to other positions, the realist can answer the following questions in particular. Why do not all our wishes come true? Why don't we succeed in everything we strive for? Because of what do scientific theories fail? (The realist: Because the world is different from what we expect, hope, assume.) Why do independent measuring methods for natural constants produce the same results? Why does it seem that such results approach a limit? Why does usually one theory among competing ones prove to be superior to all the others? (The realist

explains this convergence of research by the uniqueness of the real world examined by us.) Why is our search for invariants, e.g. natural constants, general laws of nature or conservation quantities, so successful? (In particular such invariants are for the realist indications of *objectivity* of insights, i.e., of their connection with reality and their independence of the perceiving subject.)

Now, there are many variants of realism: naïve, critical, hypothetical, scientific, convergent, internal realism. But not all of these variants are tenable. The naïve realism ("The world is how it appears to me.") is already refuted by the possibility of error, in particular by the existence of contradicting perceptions. But also the classical realistic view ("All qualities are or are not connected to the things *unaffected by interaction*, especially by observation.") is called into question by modern quantum physics. If—on the other hand—internal realism ("Real is to what a fictitious (!) conclusive description of the world successfully refers.") is a realism at all, i.e., has sufficient realistic substance, is at least doubtful.

In view of the remaining range of realisms the naturalist opts for as much realism as possible. He is a realist, because he thinks a world without man is possible, but not man (or human mind) without a real world. Space, time, matter and evolution are real to him, independent of consciousness (but perceptible by consciousness). So he opts for as much *objectivity* as possible, but for as much subjectivity as necessary.

It could seem as if this maximum-realism goes unnecessarily far beyond our minimum-metaphysics. But this is not the case: We *need* this realism, in order to *explain* everyday experiences, the course of evolution and of the sciences. Only the person who feels absolutely non need for explanation can relinquish realism.

> (iii) The method of empirical science is superior to all others when doing research in the field of nature.

The method of empirical science lives from the interplay between *theory and experience*. For a direct way from immediate experience to theory does not exist, we have to rely on *experiment and elimination of error*. Finally, all means are allowed for experimenting, i.e., to find describing, explaining and predicative hypotheses: intuition, association, analogies, crea-

tive techniques, brainstorming, dreams, visions, speculations. But because error is the rule and truth the exception, the hypotheses must undergo a *strict critique*. As far as possible, they are scrutinized by experience—with observations, measurements, and specific experiments. If then errors are discovered, one will attempt to eliminate them.

This method has turned out to be useful especially in the natural sciences; but—in the way here described—it is also applicable to all empirical sciences. Beyond that each discipline has methodological specifications that it has not in common with other disciplines, for they are tailor-made for its particular research topics.

Because of the great success of specific scientific methods, many tend to transfer them to all other disciplines. Heuristically, such an attempt is completely legitimate; but certainly it is not guaranteed that those methods can be applied universally. In this case as well—regarding methodological questions—one will learn from successful and unsuccessful experiments; in this sense the process of experiment and elimination of error is *self-applicable*.

Occasionally, naturalism is characterised by the demand that everywhere scientific methods should be applied *exclusively*. Obviously, such a *scientism* would be quite dogmatic. It would contradict the principles of critical rationalism. But naturalism as well is not tied nor has to rely on such an imperialistic attitude; perhaps it is the premise, but not the result of the scientific method, even if these results proved it to be true and so support it. In particular the *premises* of one's own action, especially that of scientists, are not found in the laboratory or by an experiment or observation. Constant analysing, critical reflection of one's own premises—of a metaphysical, methodological or moral nature—thinking in an anticipating manner in this meta-theoretical sense is not a matter of singular disciplines, but of philosophy. (That does not exclude that scientists of singular disciplines philosophise occasionally—and perhaps very successfully.)

The crucial criterion for philosophical, especially meta-theoretical positions is not their empirical verifiability (or falsifiability), but their *criticisability*. For the empirical examination of factual statements presents a particularly strict way of critique, it will be used where it really can be applied. In cases where it cannot be applied, other methods will be used as well. The superiority of the empirical method is due to the exactness of its

critical instruments; but this exactness does not establish a claim to exclusiveness.

Even if everything what is belongs to *nature* for the naturalist, man, thinking, knowledge, moral and aesthetic feeling and judgement as well, he does not take everything to be a research topic of *natural science*. This seeming paradox is based on the fact that the term 'nature' alone has another meaning as in the combination 'natural science'. Although field of natural sciences has been extended considerably by ethology, neurobiology, and a philosophy orientated towards natural science and furthermore many clear-cut borderlines have vanished, it will not be assumed that all empirical sciences, humanities and social sciences as well, have turned into natural sciences. Not only natural sciences have nature in the sense of naturalism as a topic—and this will never be the case.

But in a *hierarchy* of academic undertakings natural sciences are at the bottom and physics is the basis. It is clear for everyone that such a hierarchy exists at all. But the naturalist furthermore tries to use methods and results of lower levels of the hierarchy for a better understanding of higher ones. The question of the lowest level leads us to the next step.

(iv) Nature (world, cosmos, universe, the real) is primarily materially energetic—in temporal and causal respects as well.

An alternative (which naturalism denies) would be the assumption that the world is primarily spiritual. Between the alternative materialism–spiritualism the naturalist tends to materialism, even though not to each of its forms. The classical materialism in particular starts from the assumption that all that is real is *material*. But with Clark Maxwell physics has managed to come to realize that it makes sense to ascribe reality to *fields*, waves, and rays. If one speaks of particles at all (e.g. light particles, light quanta, photons), it concerns particles without rest mass. Such systems are not characterised by their mass, but by their energy. For that reason we use the more complex expression 'matter-energy'.

The existence of spiritual, especially mental phenomena (conditions and processes) is denied by no means. But they are held as conditions and processes of real, i.e., material-energetic systems, especially of central nervous systems of a sufficient complexity. Consequently, incorporeal

mental conditions and processes do not exist. (It could be expedient to *talk as if* such incorporeal mental phenomena existed, if the material substratum is of no importance for the problem in question.)

To speak of a *primacy* of matter-energy over other "things", especially over the mental (or the spiritual), means two things: Firstly, material-energetic systems *can* exist without mental characteristics. Secondly, mental phenomena *do not* exist without a material-energetic basis. With the observation that systems without mental characteristics *in fact* have existed we proceed to the next thesis.

> (v)   All real systems—the universe as a whole included—are subject to development, evolution, increase and decrease, rise and fall.

Modern naturalism is thus an *evolutionary naturalism.*[3] Each development can have a beginning and an end; it can—according to standards which have to be defined—go upwards or downwards. It could stagnate for some time as well; but this hardly occurs in our universe.

It is also imaginable as well that the cosmic evolution we are observing (and of which we are an interim result) is only a part of a huge *cycle* that leads back to its starting point. Our universe could—in the sense of an eternal return—go through many, perhaps an infinite number of similar or identical cycles. But there is no reason to say that this is true: We nearly do not know anything—assuming that these expressions make sense—about the time before the big bang and about the time after the (possible) final bang, the same applies to the existence, number and sort of such cycles.

Today the keyword 'evolution' is often used, almost inflationary. This extensive use leads easily to a haziness of definition, ambiguities, misuse. Occasionally 'evolution' means only *biological* evolution, then "only" biological relations, the origin of organismic species out of others, phylogenetic trees, and the factors and laws of species development are concerned. The origin of life—the biogenesis—need not be mentioned yet, not even the origin of man—the anthropogenesis. Even Darwin does not treat the

---

[3]   Thus the title of an unfortunately unrecognised book: Roy Wood Sellars, *Evolutionary Naturalism*, Chicago: Open Court 1922.

origin of life or of man in his major work *On the Origin of Species*. Also later he did not think that the time was ripe for a theory of biogenesis. His book *The Descent of Man* was published not until 1871; at this point many of his thoughts had already been anticipated by others.

But it is obvious to examine the applicability of the evolutionary idea to other systems and to extend the definition and the theory of evolution "downwards" and "upwards" as well. In our century natural and social sciences have been very successful with this attempt; it is quite right to speak of a *universal evolution* and of an evolutionary *paradigm* in an extended sense. The accuracy of concepts and laws of the biological theory of evolution has to be examined in any case, i.e., for each system and each phase of cosmic evolution, but it is not a matter of course.

Theories of self-organisation try to show more exactly—and that is their function—what characterises evolutionary processes in particular, what different evolutionary phases have in common, and what distinguishes them. The concept 'self-organisation' incorporates obviously the claim to explain the formation of complex structures and patterns "from below". So these are further steps in order to realize the naturalistic programme.

As it was expected, the downward extension of the evolutionary paradigm was easier than the upward extension. Although the origin of *life* is nowhere near being clarified, understood, and explained; but there is a wide agreement that it could happen and in fact did happen on earth and "automatically", i.e., according to the then—four billion years ago—prevailing conditions and to the laws of nature known to us, that it happened according to the naturalistic sense (*mit rechten Dingen*). The evolutionary origin of *man* as well as one among many biological species is generally acknowledged, even if unfortunately many details are unknown or unclear—regarding our enormous "personal" interest.

But it is different with the higher human faculties: Recognition, language, moral behaviour, and aesthetic judgment. In theses cases many different positions stand opposed to each other incompatibly. According to the naturalistic view the evolutionary paradigm or the explanatory approach "from below" is here not only possible and reasonable, but also successful. Ethology, socio-biology, neuro-biology, bio- and psycho-linguistics, artificial intelligence, and other disciplines bring up research results for this. These results have an impact on those philosophical disci-

plines that traditionally deal with typically human faculties: anthropology, epistemology, philosophy of language, moral philosophy, ethics, and aesthetic.

The evolutionary idea combines many academic disciplines: By suggesting that the development of different systems should be seen or classified as parts or phases of a *universal* evolution, it contributes to the unity of science (see Vollmer 1989, 41-65 and Vollmer 1995a, 59-91).

> (vi) Complex systems consists of and develop from simpler subsystems.

Evolution did not start with complex systems or a particularly complex 'super-system' that gradually decay now and lose more and more characteristics. (This idea was held for a while in regard to living creatures.) It is exactly vice versa: complex systems develop later on and have characteristics that none of the subsystems ever had. We call this appearance of new characteristics of systems *emergence* (Vollmer 1992, 183-223).[4]

If complex systems *originate* form simple ones, then it is obvious to attempt to *explain* the emergent characteristics from those of the subsystems, to *derive* the former ones from or to *reduce* them to the latter. For the naturalist who assumes the *ontic* emergence the *evolutionary argument* is most convincing argument in favour of an *epistemic* reductionism.[5] This strategy was successful, but not in all fields; that is the reason why the hindrances deserve special attention. Thus, the naturalist is near to reductionism without being bound to it indissolubly.

Can real systems be divided infinitely or does a limit of divisibility exist? There will never be a final answer to this question; for we cannot find out, if our inability to divide elementary particles any further is a matter of

---

[4] There it is discussed in detail that 'emergence' *can* be explicated in a different way; that some include especially the *non-explicability* of new characteristics "from below" as a defining feature—but we think this is awkward.

[5] Regarding the evolutionary argument, see Vollmer, G.: "Die Einheit der Wissenschaft in evolutionärer Perspektive." In: Vollmer, G. (ed.) *Was können wir wissen?* Band 2: Die Erkenntnis der Natur. Stuttgart: Hirzel 1986, ³2003, 163-199, in particular 185-189.

principle or has only practical reasons. But at the moment there is no objection to see quarks and leptons as unstructured and punctiform and so as indivisible.

(vii)   The real world is connected and quasi-continuous.

One can talk about continuity in many respects. First of all, *space* and *time* are continuous parameters that have turned out to be very useful in describing the world. The real systems we know of are interrelated, too. The reason why we do not find any completely isolated systems is not surprising at all; for they could not interact with us as observers, not even indirectly, so that we simply cannot know anything of them. Thus, one can safely assume or deny the existence of isolated objects; in none of the cases a refutation need be feared. For economical reasons the naturalist assumes a world that is connected in regard to space and time.

But the *processes* which we deal with could also proceed in an abrupt manner. In fact it seems at first glance that many discontinuities exist: quantum leaps, mutations, phase transitions, experiences of conversion, catastrophes, revolutions. In most cases it is only the resolution that matters with which a process is observed. When viewed more closely, supposedly abrupt processes prove to be more than averagely rapid, but steady as well. But it seems that this does not apply to quantum occurrences. They introduce an unsteady element in our world; with regard to that the naturalist speaks of quasi-continuity.

(viii)   Authorities that are beyond human experience are imaginable, but they are dispensable for the observation, description, explanation and interpretation of the world.

Examples for such authorities, levels, beings, powers are to be found in many myths, religions, esoteric doctrines, para-and pseudo-sciences. The existence of such transcendent authorities cannot be refuted. But that does not mean that they exist (just as their unprovability does not imply their not existing).

Does that mean that the question of existence has to be left open? Again it is—as in g—the economical principle that breaks up the symmetry: The

naturalist assumes that such authorities do not exist. Thus, he is—especially with regard to the existence of a personal god—an agnostic, or even an atheist (Vollmer 1993b, 16-31 and Vollmer 1995b, 168-184). The same applies to an afterlife.

But why should we follow such an economical or simple principle? Many academics, particularly Albert Einstein or Paul A.M. Dirac, give *aesthetical* reasons and like to talk of the *elegance*, even of the *beauty* of an economical theory. The preference for simple hypotheses to complex ones is not only a question of liking. Especially Popper stresses that *methodological* reasons as well suggest such a choice: The simpler of two hypotheses is also the one that can be more easily perused (Popper calls it: falsifiable[6]); if it is false, it is easier to be recognized as false and therefore faster to be exchanged by another. Therefore the naturalist is *firstly* a monist, atheist, determinist, physicalist, or reductionist until good arguments show that such positions are *too* simple. *Imaginable* arguments of this kind are given in the following paragraphs.

(ix)   Miracles do not exist.

What are miracles? There are two different answers to this question. It is normal to define miracles as events that infringe on the laws of nature. In this case laws of nature are regularities in the reaction of real systems. But if not all systems react in that way, as the supposed law of nature predicts, then it is *not* a strict regularity and therefore it is not a law of nature. According to that explication miracles are excluded by definition. Then the assumption that miracles do not exist is true, but only analytically.

But we understand intuitively statements about the possibility or the reality of miracles as *synthetic* statements which are true or false, not only because of linguistic reasons. Therefore we define miracles as events that break through the strict cosmic order by the presence of an extramundane authority.[7] Thus, *four* elements are crucial for this definition of miracle:

---

[6]   Karl Popper proposes to equate simplicity with the degree of falsifiability in *Logik der Forschung* (1934), Tübingen: Mohr [9]1989, section 43.

[7]   See Gordon Stein (ed.), *The encyclopedia of unbelief.* Buffalo: Prometheus 1985, entry "Miracles".

(1) The existence of a cosmic *order*.
(2) The *infringement* of the latter.
(3) The *rarity*, the exceptional character of such infringements.
(4) The active participation of an *extramundane* authority.

One could assume that the former held non-existence of transcendent powers excludes automatically miracles. But that is not quite right. Extramundane authorities could be experienced by the fact that they perform miracles perceptibly; then they would not be beyond all experience and transcendent in a strict sense.

The naturalist denies both: the existence of transcendent authorities *and* the occasional intervention of extramundane authorities in natural events. A convincing proof of miracles would refute naturalism effectively.

The fact that a naturalist rejects miracles does not mean that he would not be willing to wonder or admire natural occurrences because of their beauty, complication, functionalism. Wonderment is not only for Plato and Aristotle the beginning of philosophy and science, but also for the naturalist a valuable and typically human faculty. Natural explanation does not exclude natural experience, and rationality does not exclude emotionality.

(x) An extrasensory perception does not exist.

There might be many things that we have not discovered yet. Other channels of information unknown until now are imaginable. But for them there will be also sensory organs and measuring instruments (that have to be discovered or invented, too). But a *transmission of information without a transmission of energy* does not exist; and one can even say what minimum-energy is needed in order to transmit one bit of information in our universe (Sachsse 1971, chap. 2.4).

The naturalist faces most assumptions of parapsychology very sceptically. As far as these phenomena (which are assumed as clairvoyance, telepathy, precognition, manifestation, telekinesis, or paraphysics) are well proven at all, the naturalist will look for physical, also material-energetic powers, interactions, fields, channels of information. It does not seem that there are well proven para-phenomena up to now, even if many observations still puzzle us.

In view of unusual and unexplained phenomena, the naturalist thinks it better to take advantage of *known* laws of nature in a persistent manner. Should he fail nonetheless, so he is indeed willing to consider gaps and errors in our knowledge of nature and to look for better explanations and for *new* laws of nature as well. Scientific revolutions are in fact characterised by reconsidering even central assumptions of our theories. But a recourse to transcendence, esotericism, the extramundane, or the unnatural seems to be for the naturalist a declaration of failure. Certainly we cannot and do not have to explain everything; but *if* we want to explain, *then* the naturalist demands emphatically a restriction to natural, real, material-energetic structures. But a convincing proof of extrasensory phenomena would force the naturalist into a revision.

(xi)   Even the understanding of nature does not go beyond nature.

Understanding only succeeds by the means of our brain, i.e., a natural organ. But that such an understanding *has to* succeed is not certain; for the human brain is in the first instance only an organ for survival and therefore need not be able to recognize the world. But that an understanding *could not* succeed, has not been proved as well. Although the brain was evolutionally tested in our cognitive niche, the meso-cosmos, we have left already this meso-cosmos by the means of language and therefore we are only subjected to a few fundamental restrictions.
A naturalistic interpretation of human understanding has to assume a naturalistic solution of the problem of body and soul. Such a solution which would satisfy every position does not exist up to now. Especially concepts as meaning, intentionality, qualitative feelings ("qualia") show some difficulties (see, e.g. Strawson 1985, Papineau 1993, Keil 1993). Regarding this question naturalism is still on the *programme*.

(xii)   There is a unity of nature which could be reflected in a unity
        of science.

We have already used the idea of a *unity of nature* for some of the present theses, e.g. the primacy of the material-energetic (d), the world's character of relation (g), the rejection of transcend authorities (h). Carl Friedrich von

Weizsäcker who is—as Kant—only partly naturalist characterises his idea of the unity of nature by five presumptions (von Weizäcker 1971, 466-470):

(1) Unity of laws: A single fundamental theory applies to the whole of nature (for von Weizsäcker it is the quantum theory).

(2) Unity of objects: All natural objects consist of elementary particles that are divided only into a few classes (see iv, vi).

(3) Allness of objects: The world as a whole can be seen as one single object.

(4) Unity of experience: All experiences can be embedded in a unified space-time.

(5) Unity of man and nature: Man as a perceiving subject is a part of nature with a genetic continuity with animals and finally with inanimate systems as well (see v, vii, xi).

Obviously, von Weizsäcker does not even try to separate unity of nature from unity of science. This is regrettable; for they can be differentiated easily. Thus, it is imaginable that the idea of a unity of nature proves to be successful, but a unity of science is not achieved because of rather pragmatic reasons.

For the naturalist, the idea of the unity of nature points the way ahead. But it can be filled in different ways. A "final" formulation of this idea rather does not exist.

4. WHAT IS INDISPENSABLE FOR NATURALISM?

All these theses are meant to be *working hypotheses* that are criticisable and correctable on their part. Some of them are at least indirectly perusable; e.g. thesis d would be refuted, if forces without a material-energetic carrier were be detected, if live forms without a material basis existed, or if psychological processes without a neural (or another comparably complex material-energetic) substratum existed. Although the non-existence of such carriers cannot be proven, it can be demonstrated with much plausibility. Thus, successes and failures of natural research are in particular crucial for

a change of position. In fact one would be inclined in the past to character-
ise naturalism by a strict demand of continuity, e.g. the Leibnizian *Natura
non facit saltus*. But in view of modern physics this demand was untenable.
(It was nevertheless indirectly perusable, namely by the success of funda-
mental physical theories and their interpretations.)

The naturalist is thus willing to reconsider his demands and to change or
amplify them if necessary. Accordingly, he is methodologically near criti-
cal rationalism. (But that does not mean that all critical rationalists are or
have to be naturalists; e.g. Popper is not a naturalist—as his three world
theory shows clearly.)

Certainly, the naturalist cannot move away from each of his theses as
far as he wants to. As every conception of the world the naturalistic one
has indispensable elements: they cannot be given up without the surrender
of naturalism as a whole. This is not dogmatic, but a question of clear defi-
nition: It is certainly possible—if required—to *leave* the naturalistic posi-
tion at one's own or somebody else's wish or without a motive; but not any
position is *called* 'naturalism'. Probably, the following programmatic de-
mands are indispensable:

(1)   Only as much metaphysics as necessary! (i)
(2)   A minimum realism according to that a world without man is
        possible. (A weak version of ii)
(3)   Primacy of inanimate matter-energy (iv)
(4)   The construction of real systems from simple particles (v)
(5)   No transcendent authorities related to experience (viii)
(6)   Therefore no miracles (ix)
(7)   The mental faculties of man do not go beyond nature. (xi)

Naturalism is—as seen—still on the *programme* in regard to many issues.
The confidence that naturalist have in this programme is not so much based
on proofs—which hardly exist—as on economical principles, on the lead-
ing role in research of naturalistic theses, and on the successes. These bases
are so fundamental that it is not easy for the anti-naturalist to deprive them
of their supporting power. At the same time they show how one can argue
against naturalism: one objects to economical principles (with good rea-
sons); one shows that anti-naturalistic premises are—at least occasion-

ally—heuristically more fruitful than naturalistic ones; one does not accept actual or presumable success as an argument or denies effectively the success of the naturalistic approach. It should be scrutinized if anti-naturalistic arguments can be systematized sensibly according to that division; but of course such a scrutiny cannot be done here.

5. WHY AM I A NATURALIST?

This chapter is more personal than the previous ones. The editor has explicitly asked me to do this. There are less arguments than confessions. I think confessions are acceptable, but philosophy cannot restrict itself to it. Therefore the actual problem is not to make confessions not to suppress them, but to mark them as such, as they appear. ('Confessions' do not necessarily mean confessions of faith. I will deal with that later on.)

To characterise a position is one thing, to take up one another. I am a naturalist in the sense I have described in the previous chapters. Certainly, this attempt is based on my desire to make it clear for myself what I really mean and which philosophical "pigeon-hole" I am part of.

One can have an attitude, a belief, a conviction without thinking about the reason why. That is even the rule rather than the exception. But philosophers do reconsider; same even think that this is the main task of philosophy. Thus, one can also analyse one's own attitude: biographically, critically, by arguments. One can even ask why one is inclined to analyse things, i.e., the act of asking is employed on oneself. Then the asking will not end.

Why am I a naturalist then? I have not always been one. How and why did I become one? There is no definite moment in which it happened, no crucial experience. It was a gradual process. Often I only recognized afterwards that I did not want or was not able to agree on this or that any longer.

Naturalism as I understand it has many facets: ontological, methodological, semantic. In regard to religious questions naturalists are agnostics or atheists. Now I will report especially on the religious side.

I grew up as a Protestant. At home we said our prayers, I was baptized and confirmed, went to religious education and to church, participated in the Protestant parish youth for years, was myself leader of a youth group

and went to seminars of the Protestant Academy. Our pastor gave me a bilingual New Testament; he said I should study theology and become a curate. I was even married in church.

But I always had problems with religious *contents*. To take the bible as god's word seemed unreasonable to me. I did not like when someone answered a question with a biblical saying. I was startled as I learnt how little was known historically about the person called Jesus. Miracles already seemed implausible to me at an early stage. The problem of theodicy not only seemed unsolved to me, but insoluble. The maxim of some church scholars "Credo quia absurdum" seemed absurd to me. How should one tell the difference between absurd statements which one should believe and absurd statements which one should not believe? There was more of such doubts.

My natural scientific studies have contributed very much. The results and the methods of science as well were of importance. Science contradicts many theological statements. And it has developed methods for problem solving that are very different form the methods of theology. But the statement that results and methods are *different* does not provide a decision which results are correct and which methods are appropriate. What should or what can be done in such a case? Here philosophy comes into it.

During my philosophical studies religious or religious-philosophical questions were of no importance (anymore). Nonetheless it was effective, because I learnt and realized how important it is to be *consistent*. There are many—also and especially many scientists—that deem others things to be true on Sundays than on weekdays. It may be that one is able to see scientific and religious truths as two sides of the same reality, as "perspective", "dual", or "complementary". I do not succeed in doing this; I cannot overlook contradictions in that way.

Of course, one can attempt to avoid such contradictions. One could limit theological statements to such ones that do not ascertain claims to truth. This has been proposed occasionally; but it is obviously not that way that theology wants to take. Also one can limit oneself to fields that are not open to experience, especially to scientific experience. But how is its truth made plausible then?

It seems to me that one cannot avoid the contradictions. Accordingly, a competition between scientific and religious statements is created. And

science has developed a huge apparatus in order to eliminate such contradictions. Philosophy of science has scrutinized and extended this apparatus. As a philosopher of science I am inclined to use this tool for the competition of science and religion. It is demonstrated that many theological statements are not perusable or do not pass the examination.

But it is not the case that philosophy of science has turned me into a naturalist. I had my doubts already at an earlier stage. But it has given me the means to handle those doubts, to state them more precisely, to provide them with arguments, to substantiate them.

And still later I came to know the expression 'naturalism' for my position. Before I had only the term 'materialism' at my disposal. But it is not very apposite. Firstly, it only shows the ontological side, i.e., the material structure of the world. Secondly, this materialism has changed, even revised itself: Long since materialists accept that the world does not only consist of matter, but also of fields and energy, therefore we talk more precisely of a material-energetic structure of the world. Thirdly, many people assume that a materialist only thinks of the "material", it is all about—in simple words—money. But in this sense a materialist can be quite an idealist. Fourthly, 'materialism' implied easily dialectical and historical materialism then, and that seemed to me—in spite of much agreement—quite misleading. The term 'naturalism' came just at the right time for me; now I had a name or my position.

Should one spread one's attitude? Should one promote it? Should one attempt to convince others? In regard to this question I have a dual, but hopefully clear answer: If I am *asked*, then I advocate my view. Faust's answer to the Gretchen question "How do you feel about religion?" is—because of understandable reasons—unclear, evasive, cowardly. Faust hides the fact that he does not think the same way as Gretchen does. His god is abstract and impersonal, Gretchen's god is concrete and personal.

But I do not feel the need to convert others; *I do not do missionary work*. This attempt has already caused much disaster in world history. I am a *fallibilist*. We always make mistakes, and it may be possible that I am wrong. The fallibilist is modest: Although he has an opinion, perhaps even a strong conviction, he takes the fallibility of everyone, his own as well, into account. Perhaps he thinks that he knows what others lack; but he is

not sure about this and therefore he will not make a person do what is good for that person or even try to persuade that person.

Hardly anybody has held fallibilism more consistently than Karl Raimund Popper. He has formulated it again and again, has talked of tolerance and intellectual modesty, has recommended it strongly. But two problems remained unsolved, a theoretical and a practical one.

The *theoretical* problem is included in the question how tolerant one should or must be towards enemies of tolerance. If one is too tolerant, so it will be used, oppressed, eliminated. Therefore Popper decides in favour of the motto: "No tolerance towards the enemies of tolerance!" But this strategy is not consistent: Tolerance is undermined. In particular this strategy makes the limits unrecognizable: Given as the pretext that a person rides roughshod over tolerance, one can fight him—and consequently everybody—and that in the name of tolerance! For this problem I have no real solution.

The *practical* problem consists of the fact that—although Popper taught tolerance—he could be very intolerant as a person. I experienced that repeatedly, and many of his students, even and especially his best, suffered. He created some of his worst enemies because of that. One can excuse Popper by saying that he was not only a wise man, but also a passionate one. But the contradiction between theory and life remains—a contradiction that is called elegantly a *pragmatic* or *performative inconsistency*: One acts in another way than that one recommends. This does no harm to the legitimacy of a theory, but to its persuasiveness.

If it is true that Rousseau who has written with "Emile" a novel about education put his own children into an orphanage, then he is also guilty of such a pragmatic inconsistency. (But it is suspected that Rousseau did not really do this, but simply wanted to attract attention with this confession. But then he was not honest, and presumably this contradicts his educational ideals, too.)

Socrates taught that laws have to be obeyed, even if one thinks them unjust. According to that theory he refuses to escape in order to evade execution. Perhaps therefore he is regarded by some as the most important philosopher: He even died for his belief.

So I do not expect that all human beings, all colleagues, or all friends are or become naturalists. Of course it is agreeable to find that others share

the same opinion. Then you are spared of discussions, explanations, debates. But this is not crucial. The willingness to debate without polemics or disparagement is crucial. It is crucial that you tolerate each other.

Regarding this the fallibilist has it easy. Fallibilism is not a confession of faith. The fallibilist is willing, well, let's say should be willing to expose all views—and all confessions—to criticism: naturalism, realism, critical realism and even fallibilism as its basis. This position that also regards the critical rationalism as provisional and correctable is called *pancritical rationalism* by William Bartley (see Vollmer 1993a, 6ff. and 152f.). He is more consistent as Popper himself. Because I estimate such consistencies, I am a pancritical rationalist.

Can one be both, a naturalist and a pancritical rationalist? Or is it a contradiction, perhaps a pleonasm? I think that is not the case. Naturalism is an extensive *philosophical* position (everywhere in the world everything can be explained rationally—*überall in der Welt geht es mit rechten Dingen zu*) that is characterised by its universal claim and its demand for limitation of means. The (pan)critical rationalism is here regarded as a *methodological* position. An overlapping exists only in so far as naturalism has itself methodological elements or entails them. Essentially, the methodological tools of naturalism are that of critical naturalism. (A conversion is not allowed: Not each critical rationalist is already a naturalist. Thus, I would not call Popper a naturalist, especially regarding his three world theory and his attitude towards the problem of body and soul.)

So one can be a pancritical rationalist *and* a naturalist. And this I am consequently.

6. REFERENCES

Franzen, W. 1984: "Grenzen des Naturalismus?" *Gießener Universitätsblätter* 17, May 1984, 69-77.

Kanitscheider, B. 1988: "Probleme und Grenzen eines naturalistischen Weltverständnisses." In: Hinke-Dörnemann, U. (ed.) *Die Philosophie in der modernen Welt*. Frankfurt am Main: Lang, 603-630.

Kant, I. 1788/1968: *Über den Gebrauch teleologischer Prinzipien in der Philosophie*. Darmstadt: Wissenschaftliche Buchgesellschaft (Kant Werke, Band VIII).

Keil, G. 1993: *Kritik des Naturalismus*. Berlin: de Gruyter.

Markl, H. (ed.) 1983: *Natur und Geschichte*. München: Oldenbourg.

Nagel, E. 1955: "Naturalism Reconsidered". *Proceedings and Addresses of the American Philosophical Association*, 5-17.

Papineau, D. 1993: *Philosophical Naturalism*. Oxford: Blackwell.

Sachsse, H. 1971: *Einführung in die Kybernetik*. Braunschweig: Vieweg.

Strawson, P. F. 1985: *Scepticism and Naturalism: Some Varieties*. London: Methuen.

Vollmer, G. 1989: "Der Evolutionsbegriff als Mittel zur Synthese—Leistung und Grenzen." *Philosophia Naturalis 26*, 41-65.

Vollmer, G. 1992: "Das Ganze und seine Teile. Holismus, Emergenz, Erklärung und Reduktion." In: Deppert, W. et al. (eds.) *Wissenschaftstheorien in der Medizin*. Berlin: de Gruyter, 183-223.

Vollmer, G. 1993a: *Wissenschaftstheorie im Einsatz*. Stuttgart: Hirzel.

Vollmer, G. 1993b: "Bin ich ein Atheist?". In: Dahl, E. (ed.) *Die Lehre des Unheils*. Hamburg: Carlsen, 16-31.

Vollmer, G. 1995a: *Biophilosophie*. Stuttgart: Reclam.

Vollmer, G. 1995b: *Auf der Suche nach der Ordnung*. Stuttgart: Hirzel.

Von Weizäcker, C. F. 1971: *Die Einheit der Natur*. München: Hanser.

Wendel, H. J. 1993: "Die Grenzen des Naturalismus". In: Großheim, M. and Waschkies, H.-J. (eds.) *Rehabilitierung des Subjektiven*. Bonn: Bouvier, 85-109.

Naturalism and Theism as Competing Traditions

Nancey Murphy
Fuller Theological Seminary

1. INTRODUCTION

Some would think it strange to raise the question of the success of naturalism because in so many ways naturalism has been a merely negative position, and still remains surprisingly difficult to define positively: Should it be in terms of ontology or epistemology, and if epistemology is it explanation in terms of the physics of today or the ideal physics at the end of time? It is still *tempting*, is it not, to define it negatively, as a non-theistic or non-supernaturalistic worldview? Perhaps it could be defined epistemologically in terms of what could be comprehended by physics in the mind of an omniscient God—if only there were one.[1]

'Naturalism' is sometimes used interchangeably with 'materialism' and 'materialism' with 'physicalism'. With each of these terms there are two main sorts of theses being denied. One is theism (along with any additional supernatural beings such as angels and demons); the other is substantial dualism with regard to the person—no immaterial mind or soul. There is no reason at all to take physicalism or naturalism with regard to humans as tantamount to atheism, although surprisingly many seem to do so. As the

---

[1] The lack of clarity regarding that to which the term 'naturalism' refers can be seen by checking philosophical sources from different decades. For example, in John Passmore's *A Hundred Years of Philosophy* (1957) it is taken to refer to a small group of philosophers in the early twentieth century; compare this with the extensive list of philosophers and topics in the index of the *Routledge Encyclopedia of Philosophy* (1998).

title of this essay suggests, my concern here is strictly with the competition between theism and non-theism.[2]

Yet it is a mistake at this point in intellectual history to think of naturalism in this sense as *simply* the denial of the existence of God, rather than in terms of competing worldviews, as in a debate I heard between Richard Dawkins and Simon Conway Morris. Dawkins exclaimed that he and Conway Morris understood the science exactly the same and Dawkins kept asking Conway Morris why he insisted on adding God to it. The position that I shall take in this essay is that naturalism should be seen as a *tradition* in its own right, beginning with David Hume's corpus, and perhaps also with Baron d'Holbach's *System of Nature* (1770), which presented a systematic treatment of the world as a whole, humanity's place in it, immortality, the structure of society. The tradition includes others' accounts of the origins of religion; later Karl Marx's, Sigmund Freud's, and Friedrich Nietzsche's explanations of the persistence of religion; and a variety of later theories of ethics basing morality on human reasoning as opposed to divine will. Richard Dawkins, E. O. Wilson, Daniel Dennett, and others are current contributors to this tradition.

So, in contrast to the approaches taken by philosophers of religion for the past few centuries, the tenability of theism is not to be approached by attempting to construct relatively brief arguments for the existence of God. Instead it depends on finding criteria for rational comparison of large-scale traditions such as this relatively new naturalist tradition and one or more of the older theistic traditions. I shall draw upon resources from Alasdair MacIntyre to consider what it would take to make such a comparison.

My plan, then, will be first to present MacIntyre's somewhat technical description of a tradition and his account of how it is possible (sometimes) to make rational adjudications between competing traditions—despite the fact that each usually incorporates its own standards of rationality and accounts of truth. In part this involves examining the intellectual crises each has faced and the extent to which each has or has not been able to overcome them. Next, I look at what I take to be the most serious crises Christian theism has faced in the modern period and note its progress in addressing them. Finally, I shall describe naturalism as a developing

---

[2]   I have argued for a physicalist account of the person in my 2006.

tradition and raise questions about the crises that naturalists ought to be worrying about.

2. MACINTYRE ON THE CHARACTER AND COMPARISON OF LARGE-SCALE TRADITIONS

So far I have been using the word 'tradition' in its ordinary, non-technical sense. I now want to introduce it as a philosophical term of art, as developed by MacIntyre. Although he disclaims being an epistemologist, I have long been promoting his as the most sophisticated account of human rationality to date. However, I often find my audiences unimpressed. While this may be due to the obscurity of my writing or to the fact that his lengthy and dense books do not compress well, I also suspect that his achievement can best be appreciated against the background of the philosophy of science of the 1970s, with which many in my audiences are unfamiliar. I expect that many readers of this volume do know this history. The rationale for reading him in this light is the fact that he offered an early account of his epistemological insights in an article titled "Epistemological Crises, Dramatic Narrative, and the Philosophy of Science" (1977/1989). Here he replied to Thomas Kuhn's philosophy of science (1970) and noted shortcomings in Imre Lakatos's response to Kuhn (Lakatos 1970). I shall come back to this shortly.

The primary stimulus for further development of these epistemological insights came from his work in philosophical ethics. In *After Virtue* he argued that moral positions could not be evaluated apart from traditions of moral enquiry. Yet, without a means of showing one such tradition to be rationally superior to its competitors, moral relativism would follow (1984). In two succeeding books he has elaborated his concept of a tradition and shown by example the possibilities for such comparative judgments (1988; 1990).

Traditions generally originate with an authority of some sort, usually a text or set of texts. (Recall the role of classic texts in Kuhn's paradigms.) The tradition develops by means of successive attempts to interpret and apply the texts in new contexts. Application is essential: traditions are

socially embodied in the life stories of the individuals and communities who share them, in institutions, and social practices.[3]

(Think of Kuhn's standard experimental practices, normal science, the role of the community.) A large-scale tradition, as already mentioned, also incorporates its own theories of knowledge. For example MacIntyre points out the contrast between the epistemologies embedded in the Augustinian and Thomist traditions of the middle ages, one a Platonic epistemology altered by Augustine's doctrine of original sin; the other Aristotelian. Thus, working *within* a given tradition there will be widely agreed practices and standards for justifying claims. The difficult and more interesting question is the one addressed here: how to justify these practices and standards themselves. Finally such traditions, providing the essentials of an entire worldview, incorporate some account of ultimate reality, which sheds light on the question of the meaning of life and provides a foundation for ethics.

MacIntyre ironically characterizes Enlightenment thought as the tradition of traditionless reason. In contrast, he argues that all rationality is essentially tradition dependent. Outside of all traditions, one is morally and intellectually bankrupt. But must this *not* lead to radical relativism? Where could one stand to judge one tradition rationally superior to another? It is time to return to the relation between MacIntyre's insights and the philosophy of science.

Kuhn's *Structure of Scientific Revolutions* was criticized by many as presenting an irrationalist account of science. Lakatos responded with what he thought was a more rationalist account of scientific methodology. He argued that one could choose between competing research programs on the basis of one being more progressive than its rival (1970). Paul Feyerabend countered that this criterion is inapplicable because sometimes degenerating programs suddenly become progressive again, so one never knows when it is rational to give it up (1970, 215). I believe I am not alone in judging that Lakatos never gave a satisfactory answer to this challenge.

MacIntyre's insight is to point out that there may actually be an asymmetry between the rivals. From the point of view of one program it may be possible to explain *why* the other program failed, and failed *at just*

---

[3] MacIntyre says that his technical notion of a practice serves the same role in his philosophy as do language games in Wittgenstein's. (Conversation, May 14, 1996).

the point it did. One example is the competition between the Copernican and Ptolemaic programs. The crisis to which Galileo responded involved inconsistencies of Ptolemaic astronomy with both Platonic astronomical ideals and Aristotelian physics. The latter was inconsistent with empirical findings on terrestrial motions. Galileo resolved the crisis by reconceiving astronomy and mechanics, and in the process redefined the place of experiment in natural science. At last, the history of late medieval science could be cast into a coherent narrative. In general, MacIntyre says:

> The criterion of a successful theory is that it enables us to understand its predecessors in a newly intelligible way. It, at one and the same time, enables us to understand precisely why its predecessors have to be rejected or modified and also why, without and before its illumination, past theory could have remained credible. It introduces new standards for evaluating the past. It recasts the narrative which constitutes the continuous reconstruction of the scientific tradition. (1977/1989, 146)

What the scientific genius, such as Galileo, achieves in this transition, then, is not only a new way of understanding nature, but also and inseparably a new way of understanding the old science's way of understanding nature. The new science is taken to be more adequate than the old because it is only from the standpoint of the new science that the inadequacies of the old science can be characterized.

> It is from the standpoint of the new science that the continuities of narrative history are re-established. (1977/1989, 152)

Thus, MacIntyre claims that scientific reason turns out to be subordinate to, and intelligible only in terms of, historical reason, and criticizes Kuhn for failing to highlight these narrative connections between successive paradigms.

   MacIntyre's concern in his three volumes on philosophical ethics (1984; 1988; 1990) was to rejuvenate the Aristotelian-Thomist tradition of virtue ethics and to argue for its rational superiority to both the Enlightenment tradition and what he calls the genealogical tradition—Nietzsche and his followers. In the process he developed an account of the possibilities for rational adjudication between such large-scale traditions. The comparison depends on there being participants within the traditions with enough

empathy and imagination to understand the rival tradition's point of view in its own terms. All mature traditions face epistemological crises such as incoherence, new experience that cannot be explained, or simple inability to advance their enquiries beyond a certain point. Thus, one aspect of the adjudication between competing traditions is to construct a narrative account of each tradition: of the crises it has encountered and how it has or has not overcome them. Has it been possible to reformulate the tradition in such a way that it overcomes its crises without losing its identity? Comparison of these narratives may show that one tradition is clearly superior to another: it may become apparent that one tradition is making progress while its rival has become sterile. Echoes of Lakatos here. The important *asymmetry,* though, results when the superior tradition provides resources to characterize and explain the failings and defects of the other more adequately than the protagonists of the failing tradition are able to do.

The central claim of this essay is that the question of theism versus naturalism needs to be reformulated in terms of rival traditions. I shall try to show that naturalism is something like a MacIntyrean tradition, perhaps now with important subtraditions within it, just as there are within Christianity. I say "something like" because most adherents of naturalism do not spend their time re-interpreting and applying Hume's texts. However, within the subtraditions of Marxism and Freudianism there certainly has been this feature.

The competition for this 'tradition' cannot be understood in terms of naturalism versus theism in general, much less religion in general, but rather in terms of specific theistic or other religious traditions. So in the remainder of this essay I shall focus on modern Western Christianity.

3. CRISES IN CHRISTIANITY

In this section I shall list the intellectual crises facing modern Western Christianity that I take to be the most significant and note briefly what sorts of moves are presently being made in Christian scholarship to meet them. All would agree that the following are at least *among* the most significant. I shall list them in what I take to be their order of significance. The first is what I shall simply call the epistemological problem: How, if at all, can the Christian belief system be justified? Second and closely related

is the problem of religious pluralism: How can one claim that Christianity is true when confronted with the conflicting claims of other religions? Third, the problem of natural evil: If God is all good and all powerful, how is this to be reconciled with all of the suffering of humans and animals that is *not* caused by human misdeeds? Finally, what about the real or perceived conflicts between Christian teachings and science?

3.1. The Epistemological Problem

I believe that Princeton philosopher Jeffrey Stout has given the most incisive account of the onset of Christianity's epistemological crisis in his book *The Flight from Authority: Religion, Morality, and the Quest for Autonomy* (1981). He argues that the most significant epistemological change at the dawn of modernity involved a change in the meaning of the word 'probable.' Medieval thinkers distinguished between *scientia* and *opinio. Scientia* was a concept of knowledge modeled on geometry; *opinio* was a lesser but still respectable category of knowledge, not certain but probable. But for them 'probable' meant subject to approbation, theses approved by one or more authorities. Theological knowledge would obviously fare well in this system, being that which is approved by the highest authority of all, namely God.

However, the multiplication of authorities that occurred in conjunction with the Reformation made resort to authority a useless criterion for settling disputes. The transition to our modern sense of probable knowledge depended on recognition that the *probity* of an authority could be judged on the basis of *frequency* of past reliability. Here we see one of our modern senses of 'probability' intertwined with the medieval sense. Furthermore, if nature itself has testimony to give, then the testimony of a witness may be compared with the testimony nature has given in the past. Thus one may distinguish between internal and external facts pertaining to a witness's testimony to the occurrence of an event: external facts have to do with the witness's personal characteristics; internal facts have to do with the character of the event itself, that is, with the frequency of events of that sort. Given the "problem of many authorities" created by the Reformation, the task increasingly became one of deciding which authorities could be believed, and the new sense of probability—of

resorting to internal evidence—gradually came to predominate, making external evidence, the testimony of witnesses, count as evidence only at second remove. The transition from authority to internal evidence was complete.

Stout traces the fate of theism after this epistemological shift. The argument from design was reformulated in such a way that the order of the universe only supplies empirical evidence for God's existence, not proof, as it had in the Middle Ages. In addition, in an early stage of development it became necessary to provide evidence for the truth (that is, revealed status) of Scripture as a whole. If such evidence could be found, then the content of Scripture could be asserted as true. In a later stage it was asked why the new canons of probable reasoning should not be applied to the various contents of Scripture themselves. Here is where the challenge of higher criticism made its mark. Historian Claude Welch writes that by the beginning of the nineteenth century the question was not merely *how* theology is possible, but whether theology is possible *at all* (1972, 59). Stout's prognosis is grim: theologians must either seek some vindication for religion and theology outside of the cognitive domain or else pay the price of becoming intellectually isolated from and irrelevant to the host culture.

My own view is much less pessimistic. I have argued (1990) that theology's failure in the past to meet modern standards of justification is due not to the irrationality of theology but to the fact that modern *theories* of rationality have been too crude to do justice to theological reasoning[4] — and not only to theology, but to scientific reasoning as well. To support this latter claim, consider reactions to Kuhn's *Structure of Scientific Revolutions*. Kuhn showed that scientific practice at its best not only did not measure up to, but actually violated the methodological norms of then-

---

[4] I argue that theological schools come very close to fitting Lakatos's description of scientific research programmes. They are organized around a core thesis, generally about the nature of God. They have auxiliary hypotheses that are subject to change (doctrines), and draw upon their own sorts of evidence, some scriptural and some empirical. The empirical data include religious experience and historical events. Comparable to Lakatos's and Kuhn's theories of instrumentation, theologians and Christian practitioners have a theory of discernment to judge which putative religious experiences are genuine and thus provide legitimate data for theology.

current theories in the philosophy of science. There were two possible conclusions one could draw: either science is irrational or else the positivists' theories of rationality were inadequate. Philosophers of science have largely taken the second point of view. Only now do we have theories of human reasoning that are (in Paul Feyerabend's terms) "sly and sophisticated enough" to do justice to the complexity of scientific reasoning. I would say all the more so with regard to theological reasoning. So, I claim, the resources are now available, largely in MacIntyre's work, to solve the epistemological crisis.

3.2. Religious Pluralism

The problem of religious pluralism is one already noted by Hume in the eighteenth century, but the modern study of comparative religion did not begin until approximately 1870 (Welch 1985, 104). There have been a variety of Christian responses. One important strategy, beginning already with Friedrich Schleiermacher in 1799, is the claim that all religions are external expressions of a universal awareness of the divine. Early versions often argued that the Christian expression was superior to the others. Another strategy, begun with Max Mueller's response to the 1893 World Parliament of Religions is to claim that all of the major religions in fact have more in common than they have differences.

My own reading of the current situation is that the (supposedly) impartial study of religion has been for some time disconnected from the pursuit of Christian theology. However, there is a recent and still somewhat small resurgence of interest in giving a Christian *theological* account of religious pluralism that is both appreciative of the other religions and capable of reconciling their existence with the truth of at least the general outlines of Christian teaching. Keith Ward, recently retired Regius Professor of Divinity at Oxford, is a notable example (Ward 1994; 1996; 1998).

Despite the likely opprobrium from the academy, my own approach would be to see them as competing clusters of traditions in need of the MacIntyrean treatment that I am proposing for Christianity and naturalism. One can raise questions such as the following. Christianity has (with some difficulty, to be discussed below) managed to adapt to and incorporate the

findings of modern science. Can the other major religions do so as well? I had the privilege of attending the First International Congress on Dialogue between Science and Religion in Tehran in 2006. I believe it is too soon to tell whether Islamic traditions will be able to make this adjustment. For example, I raised the question of whether Islam could, like Christianity, incorporate a physicalist account of human nature in order to accord with contemporary biology and neuroscience. There was not a great deal of optimism expressed. If it is not possible to make such adjustments I, and many Islamic scholars as well, would count this an intellectual crisis.

3.3. Natural Evil

It has always been a part of Christian thought to consider how to understand the relation between the goodness of God and the immensity of human suffering, but it is only in the modern period that the problem of evil could be said to be of crisis proportions. Hume pointed out that if order and goodness were to be taken as evidence for a designer, then disorder and evil must be counter-evidence. The Lisbon earthquake led to mockery of G. W. F. Leibniz's famous thesis that this must be the best of all possible worlds.

There are a number of distinctions to be drawn regarding the problem of evil. There is one between the logical and evidential problems. Is the existence of evil logically inconsistent with the statements that God is omnipotent and all-good, thus falsifying traditional theism, or does it merely count as evidence against God's existence? Second, it has long been common to distinguish among three kinds of evil: first is moral evil, that is, human sin; second is natural evil, that is, suffering of humans and animals due to natural causes; and third is metaphysical evil, that is, the trying sorts of limitations to which humans are subject.

Moral evil is relatively easy to reconcile with God's goodness on the assumption that humans just *will* misuse their freedom and that freedom is a necessary condition for the kind of relation that God offers.

In earlier centuries of Christianity, natural evil was explained as a consequence of human sin. When Adam and Eve sinned, they brought upon themselves and the whole human race punishments in the form of disease and natural disasters. Not only that, their sin (or perhaps that of

angels who fell before them) disordered God's previously perfect cosmos, causing natural disasters and suffering for both humans and animals.

Now, however, with an evolutionary account of human origins and no concept of a historic fall, the question remains as to why innocent humans and animals are subject to so much suffering. The simple answer, of course, is that it comes largely from the ordinary working of the laws of nature. When children fall from trees, their bones break because of the force of gravity. Tsunamis are the result of earthquakes, and earthquakes are the result of plate techtonics.

But, one might ask, could the laws of nature not be different so that there is less suffering? Leibniz had already noted that the more we know about the world the more we realize that it is not possible to change one thing without changing others. He would thus be pleased by a recent development in science. This is what I shall call the anthropic calculations, which show that extremely slight changes in any of the constants or basic laws of physics would produce a universe unsuitable for life. Thus, it is possible now to argue that the laws of nature have to be almost exactly what they are in order for there to be life. The suffering that is caused by their operation can be seen as a necessary by-product of conditions built into creation in order that there be creatures who could respond freely and lovingly to God (see Murphy, Russell and Stoeger 2007).

The second law of thermodynamics is interesting because of its relevance to metaphysical evil. This law represents a limitation on the varieties of processes that could occur according to the other laws of physics. Thus, the effects of entropy limiting human and animal life are everywhere: the need for food; the need for clothing and shelter to conserve energy; fatigue; aging; and ultimately death. These limitations in human life are not moral evils, but certainly provide much of the motive for sin, from instances simply of being too tired to do a good deed, to robbery, murder, and many wars. Entropy plays a major role in causing suffering as well: hunger pangs, certain forms of disease, predation.

There are a number of other issues that complicate this problem, which I shall not go into here, such as the question of divine intervention to protect people from natural disasters. Nonetheless, I believe that the "necessary by-product" defense sketched here is a genuine advance in resolving this critical problem.

3.4. The Rise of Modern Science

Many conservative Christians, particularly in the U. S., and many non-religious scientists believe that science and religion are essentially incompatible, and of course this is based largely on the Galileo affair and on past and present controversies over teaching evolution in schools. However, these two cases were not *intellectual* crises. Or if they were, they were very short-lived. Even quite conservative theologians quickly found ways to integrate evolutionary theory into their accounts of creation (Livingstone 1987). I include the rise of modern science as an intellectual crisis, though, for two reasons. First, it went hand in hand with the epistemological changes detailed earlier, which constituted what I believe to have been the most severe crisis for the Christian tradition.

Second, modern physics, particularly after Newton, presented a metaphysical picture of the cosmos as a closed causal order operating on the basis of natural laws. This created a problem for understanding God's role in earthly affairs once the universe had been created. Deism was a popular option: God has no ongoing role. Liberal theologians gave up on all notions of *special* divine action—that is, miracles, answers to prayer, and so forth. Insofar as an event seems to be a special act of God, this is only because subjectively it reveals God's purposes more than the others. God's ongoing action is limited to upholding the whole natural order.

Conservative theologians object that the removal of God from history essentially guts Christianity of its meaning, and contend that the author of nature can and does intervene in the natural order. There is a lively discussion among scholars interested in the relation between Christian theology and science as to whether it is possible to give an account of special divine action without supposing that God violates the laws of nature. I believe that this is still an open question.

I do not believe that the problem of divine action, however, is itself of crisis proportions. Rather, I see the *perception* of the problem to be at the root of the development of liberal theology, and I sympathize with those for whom the liberal form of Christianity is so uninteresting as not to be worth getting out of bed for on Sunday morning. The revolution initiated by Schleiermacher was to interpret all religious language, including Scripture and doctrine, as expressions of human religious awareness. In the

hands of later and lesser thinkers, this is sometimes *all* that Christianity is about. Stout's quip regarding this type of theology: it is "giving the atheist less and less in which to disbelieve" (1981, 148). So there has been a sense in liberal Christian academia of having reached something of a dead end. This, rather than the problem of divine action itself, is the real crisis; liberal theologians have, in MacIntyre's words, found themselves unable to advance their enquiries beyond a certain point.

Enough said about the trials and tribulations of Christian scholarship. This should be enough to illustrate an important claim that MacIntyre makes against relativists. Relativists are likely to assume that proponents of one tradition will always see problems with rival traditions but be blind to problems with their own. This is certainly not the case, and many serious thinkers have judged one or more of these crises to be irresolvable and have rejected the tradition as a whole.

4. THE BIRTH OF THE MODERN NATURALIST TRADITION

James Turner makes a startling claim in his highly regarded book *Without God, Without Creed: The Origins of Unbelief in America* (1985). He argues that disbelief was *not* a live option in the U. S. until roughly between 1865 and 1890. This is surprising because we are all aware of proofs for the existence of God going back through the Middle Ages to ancient Greek philosophers. I shall not comment on the ideas available in the ancient period; however, it has recently become common to see medieval philosophers and theologians as *not* intending to persuade atheists to believe in God — since there were none — but rather as engaging in the much more modest task of showing that reason could justify belief in a God already accepted on other grounds and for other reasons. The so-called medieval synthesis made God so central to all branches of knowledge and all spheres of culture that it was *in*conceivable that God does *not* exist.

I have already described the difficulties created for theologians by the rejection of authority as a proper epistemological category. The irony is that the change can be traced back to Christians themselves for not being able to settle their differences after the Reformation. The source of agnosticism can also be traced to the Reformation. If one thinks of the

agnostic not as one who simply has not formed a judgment on the existence of God, but rather as one who has concluded that human reason is incapable of making such a judgment, the story traces back to Catholic apologists in the Renaissance such as Michel de Montaigne. These apologists revived ancient skeptical methods to show that there is no rational way to decide between Protestant and Catholic claims. Therefore the only sensible course of action is to stay within the established (that is, Catholic) faith. The availability of these skeptical arguments helped pave the way for atheism, of course: if one cannot tell whether the Protestant or Catholic or Jewish version is correct, then maybe none is (Popkin 2003, ch. 3). But a variety of other factors were needed to justify a positive rejection of religious belief.

Philosopher Merold Westphal helpfully distinguishes two sorts of atheism (1993). One he calls evidential atheism, well represented by Bertrand Russell's account of what he'd say if he were to meet God and God asked why he had not been a believer: Not enough evidence God! Not enough evidence! Given the difficulties already noted in adapting theological reasoning to modern canons of rationality, this response is readily understandable.

But if religious claims are false then one needs an *explanation* of why they are so widely believed; just as, if there are no witches, we want to know what caused people to believe there were. David Hume in Britain and Baron d'Holbach in France in the eighteenth century began the attempt to explain the origin of religion naturalistically. They argued that religion is a response to fear of the unknown, coupled with superstitious attempts to control or propitiate unseen powers. Such attempts continue today, as I shall note later.

But why does religion persist in the modern world, now that we understand natural causes? The explanations here come from Westphal's second variety of atheists, the masters of suspicion. Marx, Nietzsche, and Freud practice the hermeneutics of suspicion, the

> attempt to expose the self-deceptions involved in hiding our actual operative motives from ourselves, individually or collectively, in order not to notice (. . .) how much our beliefs are shaped by values we profess to disown. (Westphal 1993, 13)

These three develop their suspicion with primary emphasis, respectively, on political economics, bourgeois morality, and psycho-sexual development, but each also subjects the religion of Christendom to devastating critique.

Two further steps were needed to make atheism a truly viable position. It would be possible to say that religion may be an illusion, but a harmless or even beneficial illusion in that it shores up morality. So two sorts of arguments were needed. One sort was to show that religion did not serve to reveal anything about the moral order that we could not get just as well by the use of human reason. Most of the work in philosophical ethics during the modern period had this as its aim. The other was to adduce historical evidence to the effect that religion has, in fact, promoted the worst evils in history—or at least more evil than good.

So within the space of two and a half centuries, roughly from 1650 to 1890, unbelief has become a live possibility. But, as I said at the beginning, this is not merely the excision of God from an otherwise common worldview, but rather the slow development of a *rival* tradition alongside the various theistic traditions and subtraditions.

Recall that a tradition, as I am using the word, is essentially a worldview, thought of in terms of its historical development. As such, it incorporates an account of ultimate reality and an account of what is most important in human life. The latter is essential as a foundation for ethics. It also involves an epistemology. A tradition is socially embodied in social practices and institutions. Let us consider some of these practices.

One might not think of the discipline of history as a naturalist practice, but one of Hume's chief philosophical and historical goals was to supplant the traditional Christian story line of creation, fall, and redemption by a new unity of action based along secular and humanistic lines. His six volume *History of England* was written from a purely secular point of view in order to show that history can be understood perfectly well without the "prophetic-providential" mode of interpretation that was common in his day (Livingston 1984). Now even Christian historians practice their craft on the basis of naturalist assumptions.

It is probably fair to say that the most important practices and institutions embodying the naturalist worldview are found in science. After the demise of the physico-theologies of the seventeenth century, the natural

sciences began to be distinguished from natural theology. Amos Funkenstein credits Immanuel Kant with the most systematic and complex endeavor "to emancipate science from its theological baggage" (1986, 346).

These are practices parallel to those of Christian scholars. There are also now secular versions of practices that used to belong solely to the church, such as marriage by a justice of the peace. A legal system has been developed that is independent of canon law. Psychotherapy competes with spiritual direction.

The naturalist account of ultimate reality, of course, is the universe itself. It is interesting that some naturalists give this thesis a religious tone and salvific trappings. For example, Carl Sagan offers a peculiar mix of science and what can only be called 'naturalistic religion'. He begins with biology and cosmology but then uses concepts drawn from science to fill in what are essentially religious categories—categories, by the way, that fall into a pattern surprisingly isomorphic with the Christian conceptual scheme. He has a concept of ultimate reality: The Universe is all that is or ever was or ever will be. He has an account of ultimate origins: Evolution with a capital E. He has an account of the origin of sin: the primitive reptilian structure in the brain, which is responsible for territoriality, sex drive, and aggression. His account of salvation is gnostic in character — that is, it assumes that salvation comes from knowledge. The knowledge in question is scientific knowledge, perhaps advanced by contact with extra-terrestrial life forms who are more advanced than we. Sagan's account of ethics is based on the worry that the human race will destroy itself. So the telos of human life is simply survival. Morality consists in overcoming our tendencies to see others as outsiders; knowledge of our intrinsic relatedness as natural beings (we are all made of the same star dust) can overcome our reptilian characteristics (Ross 1985).

Richard Dawkins offers a naturalistic account of the meaning of life: he believes in a universe indifferent to human preoccupations, one in which the good life involves pursuing all sorts of closer, warmer, human ambitions and perceptions," including especially "the feeling and awed wonder that science can give us. . . ."

This is

> one of the highest experiences of which the human psyche is capable. It is a deep aesthetic passion to rank with the finest that music and poetry can deliver. It is truly one of the things that makes life worth living and it does so, if anything, more effectively if it convinces us that the time we have for living is finite. (1998, x)

Mary Midgley's book, *Science as Salvation* (1992) provides an extended argument and set of examples to support the claim that naturalism is more than a philosophical position allied with the sciences themselves, but is rather a worldview and a way of life, with its own mythology and ultimate values.

Now, if I have made my point that we think of naturalism as a something like a MacIntyrean tradition, can we show that it is facing any major crises? I shall propose two possibilities: one is the explanation of the persistence of religion now that we are all supposed to know of its primitive origins and the disguised motives that have kept us in its thrall. The other is providing an adequate account of the moral binding-ness, to coin a term, of morality.

4.1. Explaining Religion's Persistence

I have claimed that two necessary tasks for the naturalist tradition have been to explain the origin of religion and also its persistence now that the ignorance upon which its origins are presumed to be based has been surpassed. Hume and others offered explanations of the origin of religion from the beginning of the naturalist tradition, and I am not sure that there have been any improvements here. The work of Pascal Boyer is currently receiving considerable attention. One of Boyer's theses is that religious concepts are informed by very general assumptions from what he calls domain concepts such as person, living thing, man-made object. However, they violate usual inferences regarding these concepts. A spirit or ghost is a special kind of person that is disembodied and can go through walls. The generality of these domain specific concepts explains why similar religious concepts appear in widely different cultures (2003).

For Comte and Marx, religion was supposed to fade away as society developed. So I believe that the interesting question is whether there is an

adequate explanation to be found for the persistence of religion. And there is a self-referential twist here. I begin my education as an intelligent believer (if that is not an oxymoron). I attend college and am exposed to the causal explanations for my belief. Should that not make me as suspicious of my own belief as the masters of suspicion are themselves of the beliefs of others? Should I not then examine my own motives, and might I not find that, say, Freud is right? Yet one of the most sympathetic treatments of Freud, Marx, and Nietzsche is by Westphal, an evangelical Protestant, in a book that he wrote *for Christians,* and suggested *as a Lenten meditation!*

There are certainly explanations of religion to be found today that are by comparison to the old masters, less sophisticated: Dawkins's, for example (1998). Freud's explanation is based on his estimation of human life as constant conflict between the individual's most powerful drives and the worlds of both nature and culture. Culture demands restrictions on impulses, and nature ultimately destroys us through sickness, aging, and death. Religious doctrines are illusions, beliefs induced because they fulfill deep-seated desires. The desire is for an all-powerful and benevolent father who will compensate us in another life for the permanent internal unhappiness that we experience in this one.

According to Dawkins, religious belief held into adulthood is a function of the person not getting over the necessary gullibility of children that allows them to be apt learners, combined with the tendency children have to retain rigidly the lessons that have been drilled into them.

Freud and Dawkins are both influenced by James Frazer's thesis about the origins of religious ceremonies. Frazer traces them to symbolic or representational thinking wherein causal connections are expected between things that resemble one another. While Dawkins merely repeats Frazer's thesis, Freud's very complex theory involves insights into the way believers acknowledge small sins and atone for them as a means of hiding from themselves their deep and total sinfulness.

Two significant new moves in the tradition are attempts to explain religion neurobiologically and by means of cognitive science. The first I see as highly questionable. The second requires much more serious evaluation.

Brain scans of subjects in prayer have received a great deal of press, but make the mistake of confusing one very unusual sort of religious experience with the whole of religion. There was a fine article on neurotheology by Sharon Begley in *Newsweek* magazine several years ago (2001). It was followed by a critical response by Kenneth Woodward, whose main point is that it is a mistake to confuse spiritual experiences of any sort with religion (2001): "Losing one's self in prayer," he says,

> may feel (. . .) uplifting, but these emotions have nothing to do with how well we communicate with God. In fact, (. . .) the sense that God is absent is no less valid than the experience of divine presence. The sheer struggle to pray may be more authentic than the occasional feeling that God is close. (. . .) Neurotheologians also confuse spirituality with religion. But doing the will of God (. . .) involves much more than prayer and meditation. To see Christ in the person of an AIDS victim or to really love one's enemy does not necessitate a special alteration in the circuits of the brain. (. . .) In short, religion comprehends a whole range of acts and insights that acknowledge a transcendent order without requiring a transcendent experience.

However, if the neurotheologians have too narrow a concept of religion or religious experience, so too have theologians themselves for the past century or more, along with spiritual writers and earlier scientific students of religion. In contrast, a number of contemporary writers, such as Catholic theologian Nicholas Lash, want to deny that there is *any particular* "division of life" which is the privileged place of our encounter with God. Religion that is "something real" arises whenever humans have God in mind with the might of their being (1986).

Another new movement in the study of religion is the application of the tools of cognitive science. An impressive contributor here is Harvey Whitehouse (2004). Much of the study of religion has taken an implicit approach, that is, attempting to explain religious representations in terms of various functions that the devotees themselves would not recognize. Whitehouse sees some value in this sort of work, but emphasizes that the explicit content of religious ideas must be taken into account as well. Religious representations need to be understood in part on the basis of universal cognitive "hardware" but also on the basis of the mechanisms

that account for the selective spread and retention of religious representations.

Whitehouse has worked out a clear distinction between what he calls imagistic and doctrinal forms of religion. The imagistic clearly appeared first in human history; it is found among small cohesive groups and is characterized by infrequently repeated and highly emotionally arousing rituals. Doctrinal religion is usually practiced by larger, less intensely related groups, and involves complex semantic schemas. The hold of imagistic religion on its members comes largely from the huge personal cost of the initiations members have gone through and the "revelations" they have had in attempting to find meaning in that experience. Persistence in the doctrinal mode depends initially on rhetorically compelling prophets or preachers, but to be maintained must be reinforced by repetition that is neither too lax nor too oppressive.

It is not possible here to convey the detail and subtlety of Whitehouse's work or that of others in the cognitivist school. It is certainly the case that this research program needs to be taken seriously before it would be possible to claim that the persistence of religion cannot be explained naturalistically.

4.2. Ethics

If the modern naturalist tradition began with Hume and his arguments against the necessity of postulating God to uphold morality, then we may be coming full circle. Two of the most respected philosophical ethicists of this generation have concluded that modern moral reasoning is in a state of disorder and that the disorder could be mended by returning to a theistic justification. In a thin volume with the modest title *Morality*, Bernard Williams surveys the major approaches to ethics from Antiquity to the present (1972). He finds most of them defective in that they are not capable of answering the question why be moral (at all)? However, there is also a sort of theory

> that (. . .) seeks to provide, in terms of the transcendental framework, something that man is for: if he understands properly his role in the basic scheme of things, he will see that there are some particular sorts of ends which are properly his and which he ought to realize. One archetypal form of such a view is the belief that

man was created by a God who also has certain expectations of him (Williams 1972, 63).

However, he says, it has been practically a philosopher's platitude that even if a God did exist, this would make no difference to the situation of morality. But Williams believes this platitude to be based on mistaken reasoning:

> If God existed, there might well be special, and acceptable, reasons for subscribing to morality. (Williams 1972, 72)

Unfortunately, concludes Williams the atheist, the very concept of God is incoherent; religion itself is incurably unintelligible.

MacIntyre has taken very seriously the challenge of Nietzsche's critique of traditional morality, but finds little in modern thought with which to counter it. The development of theories in philosophical ethics from Hobbes at the beginning to the Bloomsbury group in the early twentieth century has been a failed attempt to provide a theoretical rationale for traditional morality. This has led him to conclude that modern moral discourse is in a grave state of disorder. He makes a pointed analogy: contemporary moral discourse is comparable to a simulacrum of science after a know-nothing regime has killed the scientists, burned the books and trashed the laboratories. Later, fragments of scientific texts are read and memorized, but there is no longer any recognition of the *point* of science (1984, 1f.).

Similarly, MacIntyre says, our moral language is a hold-over from the past, but we have forgotten the original point of morality. In particular we have forgotten the context that once gave it its meaning. What we moderns (and postmoderns) have lost is any notion of the ultimate purpose or *telos* of human life. Such accounts of the human *telos* used to be provided by traditions, usually religious traditions, but sometimes, as in Aristotle's case, by a metaphysical tradition. MacIntyre argues that the correct form of ethical claims is something like the following, conditional statement: "If you are to achieve your *telos*, then you ought to do x." It is a peculiar feature of modern Enlightenment views of ethics that their proper form has been taken to be apodictic: simply, "you ought to do x." Modern philosophers have developed competing theories regarding the most basic

moral claims: you ought to act so as to achieve the greatest good for the greatest number; versus: you ought to act so that the maxim of your action can be willed universally. But because morality is taken to be autonomous—that is, unrelated to other knowledge—there is no way to arbitrate between these most basic construals of the moral 'ought.' This impossibility results in the interminability of moral debates in our society. However, the interminability should not, says MacIntyre, be taken as the intrinsic nature of moral discourse, but ought rather to be seen as a sign that the entire Enlightenment project has taken a wrong turn.

If MacIntyre is correct in his claim that the original meaning of the moral 'ought' has been lost, it is not surprising that most modern moral theories have attempted to reduce morality to something else: pleasure, enlightened self-interest, sympathy, social convention. Of course, emotivism is the most radical reduction: moral judgments merely express one's attitudes or feelings toward an action or state of affairs.

The most recent attempts to account for ethics aim to reduce ethics to biology. E. O. Wilson says:

> Self-knowledge is constrained and shaped by the emotional control centers in the hypothalamus and limbic system of the brain. These centers flood our consciousness with all the emotions—hate, love, guilt, fear, and others—that are consulted by ethical philosophers who wish to intuit the standards of good and evil. What, we are then compelled to ask, made the hypothalamus and limbic system? They evolved by natural selection. That simple biological statement must be pursued to explain ethics and ethical philosophers. (1975, 153)

Michael Ruse presents a more sophisticated argument for evolutionary ethics than many of his predecessors (1986). He recognizes, as some apparently do not, the difference between 'altruism' as a moral term and 'altruism' as it is used in biology to describe animal behavior that contributes to the survival of the group. I shall use 'altruism[m]' for the moral concept; 'altruism[b]' for the biological concept. Ruse suggests that whereas insects and lower animals are genetically programmed for altruism[b], humans have instead been selected for a disposition toward altruism[m]. Thus, he is able to argue for an evolutionary source for altruism[m] without confusing it with altruism[b].

However, having properly distinguished moral behavior from superficially similar animal behavior, he then goes on to argue that morality, thus properly understood, has no possible rational justification:

> The evolutionist is no longer attempting to derive morality from factual foundations. His/her claim now is that there are no foundations of any sort from which to derive morality—be these foundations evolution, God's will or whatever. (Ruse 1986, 234)

Since there can be no rational justification for objective moral claims, what is needed instead is a causal account of why we believe in an objective moral order. Ruse's answer is that the survival value of altruism[m] does in fact provide such an explanation.

In particular, the evolutionist argues that, thanks to our science we see that claims like "you ought to maximize personal liberty" are no more than subjective expressions, impressed upon our thinking because of their adaptive value. In other words, we see that morality has no philosophically objective foundation. It is just an illusion, fobbed off on us to promote altruism[b]. So Ruse's account, while more sophisticated than Wilson's in that he fully appreciates the conceptual difference between morality, on the one hand, and sentiment, convention, etc. on the other, is most starkly reductive: moral objectivity is merely an illusion.

The lack of moral objectivity may seem *not* to be a problem so long as we all agree on the basic outlines of morality, such as the idea that altruism is a good thing. But so far there has been no answer to Nietzsche, an atheist looking at the same Darwinian biology as the other naturalists. He regards other-regarding, benevolent, justice-seeking, self-sacrificial 'morality' as "slave morality." Christians and others of their kind advocate it because they are usually weak and oppressed, so requiring justice from the rich and powerful is in their self-interest. It was people such as these who invented the distinction between good and *evil* so that they, in their resentment, would have a pejorative term for those who reject their slave morality. Having the label of "evil" for these others feeds the masses' sense of moral superiority. Nietzsche writes:

> From the beginning Christian faith has been sacrifice: sacrifice of all freedom, of all pride, of all self-confidence of the spirit; it is simultaneously enslavement and self-derision, self-mutilation. (1886/2002, 44)

For his part, the herd man of today's Europe gives himself the appearance of being the only permissible type of man and glorifies those characteristics that make him tame, easy-going and useful to the herd as the true human virtues, namely: public spirit, goodwill, consideration, industry, moderation, modesty, clemency, and pity" (1886/2002, 86f.).

There is scholarly debate about the extent to which Nietzsche's ideas influenced the rise and acceptance of Nazism and the eugenics movement. But apart from any actual historical exemplification, we can certainly see how different a Nietzschean world would be from one based on the mild-mannered altruism that Ruse, Wilson, and others assume that biology favors.

So I conclude that the lack of an account of the moral ought, recognized as a feature of the modern naturalist tradition by some of its most sophisticated proponents (MacIntyre was himself one of these) represents a severe crisis for the tradition.

5. CONCLUSION

It is time to sum up. Of course, a MacIntyrean evaluation of two rival traditions is not possible in one short essay. I shall only say one thing in favor of the theistic point of view, which I may have appeared to do more to discredit than to support. It could easily be argued that the cognitivist approach to the spread and persistence of religion could be *incorporated into* a theistic tradition, and in fact used to good effect to improve preaching, church attendance, and so on. Also, if MacIntyre and Bernard Williams are correct, a theological account of reality solves the problems of the foundations of morality. This would be one small step in arguing for a theistic tradition.

However, I shall be content if I have done nothing more in this essay than to have changed the terms of the debate. No more arguments for the existence of God, or arguments against the existence of God, but rather,

consideration of what it would mean to show that a naturalist tradition or a theistic tradition is rationally superior to its most significant rivals.[*]

6. REFERENCES

Begley, S. 2001: "Religion and the Brain." *Newsweek,* May 7, 52-57.

Boyer, P. 2003: "Religious Thought and Behaviour as By-Products of Brain Function." *Trends in Cognitive Sciences* 7, 119-124.

Dawkins, R. 1998: *Unweaving the Rainbow: Science, Delusion and the Appetite for Wonder.* Boston and New York: Houghton Mifflin.

Feyerabend, P. 1970: "Consolations for the Specialist." In: Lakatos, I. and Musgrave, A. 1970, 197-230.

Funkenstein, A. 1986: *Theology and the Scientific Imagination from the Middle Ages to the Seventeenth Century.* Princeton, NJ: Princeton University Press.

Kuhn, T. 1970: *The Structure of Scientific Revolutions.* 2nd ed., Chicago: University of Chicago Press.

Lakatos, I. 1970: "Falsification and the Methodology of Scientific Research Programmes." In: Lakatos, I. and Musgrave, A. (eds.) 1970, 91-196.

Lakatos, I. and Musgrave, A. (eds.) 1970: *Criticism and the Growth of Knowledge.* Cambridge: Cambridge University Press.

Lash, N. 1986: *Easter in Ordinary: Reflections on Human Experience and the Knowledge of God,* Charlottesville. VA: University Press of Virginia.

Livingston, D. 1984: *Hume's Philosophy of Common Life.* Chicago: University of Chicago Press.

Livingstone, D. 1987: *Darwin's Forgotten Defenders: The Encounter Between Evangelical Theology and Evolutionary Thought.* Grand Rapids, MI: Eerdmans.

MacIntyre, A. 1977/1989: "Epistemological Crises, Dramatic Narrative, and the Philosophy of Science." *The Monist* 60, 453-472. Reprinted in

---

Hauerwas, S. and Jones, L. G. (eds.) *Why Narrative? Readings in Narrative Theology.* Grand Rapids, MI: Eerdmans, 138-157.

MacIntyre, A. 1984: *After Virtue* (2nd ed.). Notre Dame: University of Notre Dame Press.

MacIntyre, A. 1988: *Whose Justice? Which Rationality?* Notre Dame: University of Notre Dame Press.

MacIntyre, A. 1990: *Three Rival Versions of Moral Enquiry: Encyclopaedia, Genealogy, and Tradition.* Notre Dame: University of Notre Dame Press.

Midgley, M. 1992: *Science as Salvation: A Modern Myth and its Meaning.* London, New York: Routledge.

Murphy, N. 1990: *Theology in the Age of Scientific Reasoning.* Ithaca, NY: Cornell University Press.

Murphy, N. 1998: *Anglo-American Postmodernity: Philosophical Perspectives on Science, Religion, and Ethics.* Boulder, CO: Westview Press.

Murphy, N. 2006: *Bodies and Souls, Or Spirited Bodies?* Cambridge: Cambridge University Press.

Murphy, N, Russell, R. J., and Stoeger, W. (eds.) 2007: *Physics and Cosmology: Scientific Perspectives on the Problem of Natural Evil.* Vatican City State: Vatican Observatory.

Nietzsche, F. 1886/2002: *Beyond Good and Evil.* Edited by Horstmann, R-P. and Norman, J. Cambridge: Cambridge University Press.

Passmore, J. 1957: *A Hundred Years of Philosophy.* London: Duckworth.

Popkin, R. 2003: *The History of Scepticism From Savanarola to Bayle,* rev. ed. Oxford: Oxford University Press.

Ross, T. 1985: ''The Implicit Theology of Carl Sagan.'' *Pacific Theological Review* 18, 24-32.

Ruse, M. 1986: "Evolutionary Ethics: A Phoenix Arisen." In: Thompson, P. (ed.) *Issues in Evolutionary Ethics.* Albany: State University of New York Press, 225-247.

Stout, J. 1981: *The Flight from Authority: Religion, Morality, and the Quest for Autonomy.* Notre Dame: University of Notre Dame Press.

Turner, J. 1985: *Without God, Without Creed: The Origins of Unbelief in America.* Baltimore, MD: Johns Hopkins University Press.

Ward, K. 1994: *Religion and Revelation.* Oxford: Clarendon.

Ward, K. 1996: *Religion and Creation.* Oxford: Clarendon.

Ward, K. 1998: *Religion and Human Nature.* Oxford: Clarendon.

Welch, C. 1972: *Protestant Thought in the Nineteenth Century: Volume 1, 1799-1870.* New Haven: Yale University Press.

Welch, C. 1985: *Protestant Thought in the Nineteenth Century: Volume 2, 1870-1914.* New Haven: Yale University Press.

Westphal, M. 1993: *Suspicion and Faith: The Religious Uses of Modern Atheism.* Grand Rapids, MI: Eerdmans.

Whitehouse, H. 2004: *Modes of Religiosity: A Cognitive Theory of Religious Transmission.* Oxford: Altamira Press.

Williams, B. 1972: *Morality: An Introduction to Ethics.* Cambridge: Cambridge University Press.

Wilson, E. O. 1975: *Sociobiology: The New Synthesis.* Harvard: Harvard University Press.

Woodward, K. 2001: ''Faith is More than a Feeling.'' *Newsweek,* May 7, 58.

How Successful Is Naturalism?
Talking about Achievements beyond Theism and Scientism

Thomas Sukopp
Technical University of Braunschweig

To state it up front: I am a naturalist. I do not think that this statement is more than a confession at the moment. Confessions are legitimate, but of course they should not supersede reasons. I therefore should have good arguments in which respects naturalism *is* successful or at least *might be* successful. My task is threefold: I have to explain what I mean by using the slippery word "naturalism". I will also argue that we disagree not only because of fruitless logomachy. I will define naturalism as the—hopefully fruitful—middle between a narrow-minded and self-destructing scientism on the one hand and a tradition in philosophy that is primarily occupied with its own history on the other (1). I deny the thesis that we should not spend much time on explications. Some of my opponents like Nancey Murphy or Michael Rea (both in this volume) define naturalism in such ways that are not acceptable. I will return to this point. To put it metaphorically: Naturalism is neither an impasse, nor a one-way street or a highway.

By arguing against Quine I will try to attack a position that results in inconsistencies or is even self-refuting. This section leads to challenging problems naturalism should deal with, e. g. normativity or the limits of scientific explanations (2).

The third and final chapter sketches answers to the question of this volume: "How successful is naturalism?" I will give examples for success as well as for serious problems that indicate where naturalism is not successful (3). This tour-de-force-ride is inevitably abridging.[1]

---

[1] I should say that in my point of view the two main fields of research about naturalism and non-naturalism are: Metaphilosophical implications of naturalism/

1. WHAT IS NATURALISM?

"Tell me, what do you think about naturalism?" Whether one is a naturalist or not is no "Gretchen-Question" (see *Faust I* by Goethe). I therefore do not twirl like Faust did when Gretchen asked him. To say "Sure, I am a naturalist" should not be the end of a discussion but the beginning of one. This label neither means that someone has had a "philosophical coming out" nor that we of course agree with some naturalistic "wisdoms" which are—taken as serious theses—more than platitudes. A "single naturalism" does not exist. I found at least 30 different terms of "naturalism". Are these terms merely different designations? No, things are different.

I am not very much interested in doing philosophical taxonomy, but it has to be done. First, naturalism is not only, but also, a research program *and* can be identified by various theses. I only can give some hints why we should be careful with naturalistic labels. On the one hand, naturalism and their opponents spend too much time building straw men (see e.g one famous dictum of Sellars). On the other hand, naturalism tends to be an ideological point of view, the German term "Weltanschauung" fits better (see e.g. Keil & Schnädelbach 2000).

1.1. Naturalism: Levels

I—like many others—distinct between ontological, methodological and epistemological (in a narrow sense) naturalism. We might add metaphilosophical or analytical naturalism. Here I leave these and other subtleties aside. Nevertheless, this classification is plausible. Ontology, epistemology and methodology raise questions of different complexity. And they interdepend in a multifactorial mode. Ontology is the most basic discipline because it asks what exists at all. (Therefore ontological austereness—"anaemic ontology"—as a naturalistic attitude is out of place.) What do we know about the existing entities? That's one central epistemological question. Finally methodologists ask for strategies which lead to knowledge.

---

non-naturalism and "weltanschauliche", e.g. ideological, consequences, entirely based on anthropological views.

Of course, things are not that simple. The relationship between ontological, epistemological and methodological naturalism is no simple relation, since epistemology includes ontology and so on. An exponent of a specific ontology, e.g. physicalism, will certainly debar some epistemological questions, but its epistemology does not completely result from its ontology. A defender of eliminative materialism does not ask for the effectiveness of "mental things", because they do not exist. After these longer preliminaries let us classify naturalism. At the end of chapter 1, I will outline naturalism as a research program. I hope that my disambiguation will partly clear up "the lack of clarity" which Nancey Murphy (in this volume footnote 1) notices in view of common notions of "naturalism".

1.1.1. Ontological naturalism

A *weak ontological naturalism* is compatible with supernaturalism. It does not explicitly eliminate any possibility of a "higher" realm, of fundamental other nature and habit, beyond our—by (natural) laws accessible—world. This realm may be called a transcendent sphere over or beyond the world. Our supernaturalist, on the other hand, does not expressly assume such a realm. Consequently his supernaturalism is, in the last analysis, a world-immanent naturalism. This ontological "naturalism" is too weak to deserve the label "naturalism". (Of course I do not suggest to affix a seal of "proven quality" on a naturalistic position simply because we name it this way.) Who denies such a weak "naturalism", since it claims too much? I guess not too many philosophers.

Strong naturalism asserts that the distinction between nature and a realm over or beyond nature is preposterous. "World", "cosmos" or "universe" include every actually existing "thing". There is no place (and space) for supernatural entities. This strong naturalism is a justifiable, partly justified position that even can be tested when it raises empirical claims. It surely has metaphysical elements, e.g. referring to "world", which is often, but not necessarily, combined with realistic positions. This naturalism is no arbitrary supposition but rather follows from methodological principles of science. In consequence of these well-known and widely accepted principles, hypotheses and theories should be testable, or—for typical philosophical issues—criticizable.

Whatever it is we would like to test, we have to interact directly or indirectly. We can, in turn, only interact with systems which operate according to (natural) laws (see e.g. Bunge &, Mahner 2004).

Following this strong ontological naturalism, long lists tell you what does not exist: God, Gods, human or other souls as pure substrates, angels, demons, ghosts, miracles, telepathy or other typical transcend entities.[2] For a short positive list of what exists, see e.g. Quine, who offers a very short list: Matter and classes. This is in fact an "ontological desert" (Koskinen 2004) and even the naturalist, including the author of this paper, thirsts for some ontological commitment. Other naturalistic ontologies are richer (cf. e.g. Bunge & Mahner 2004).

1.1.2. Epistemological naturalism

Constitution, awareness and justification of knowledge about the world is a process that can be developed by science (and humanities and other intellectual efforts; see e.g. Hedrich 1998, 26, or Kornblith 2002, Goldman 1994a and 1994b). Every defender of this position should also hold the view that we can identify a "package deal" consisting of ontological naturalism and moderate methodological naturalism. The latter insists that if knowledge-generating processes are part of the one and only (natural) world, and if they are gained partly by scientific methods, then we trust epistemologically that science sometimes helps us to answer epistemological questions. I will return to this point in Chapter 3.

1.1.3. Methodological naturalism

To cut a long story short: Methodological naturalists (see e.g. Philosophers like Goldman, Kitcher, Koppelberg, Kornblith, Laudan, Quine or Vollmer) rarely claim that science (i.e., natural sciences like physics, chemistry and biology) only and exclusively govern our methods in a form of an "autarchy" (against this view see e.g. Rea 2007, in this volume). Methodological naturalism should not be seen as tantamount to scientism. (See e.g. Goldman 2006 about the social dimensions of research and

---

[2] This list includes not only entities. I do not intend to mix or confound categories. A generous reading of "miracles" in this respect could be: An entity that is indispensable for the existence of "miracles" is e.g. a Christian God.

acquiring knowledge). Methodological naturalism is nevertheless not "harmless" (Almeder 1998), because it emphasises the continuity of philosophy with science (and humanities). We could e.g. see human reason as a fallible and empirically criticisable capability.

After this very brief sketch, let me outline an approach that is quite helpful in distinguishing between moderate and radical methodological naturalism. Dirk Koppelberg (2000, S. 82f.) formulates this species of naturalism as a package of three theses. The first one suggests that philosophy is no adequate foundation of sciences. It rejects any "Prima Philosophia". Secondly, philosophy "has no epistemically privileged point of view compared with science (and humanities)" (Koppelberg, see above). Finally, the third thesis formulates and demands the application of scientific investigations and results within philosophy. I will try to show in Chapter 3 that these claims are justified.

In order to distinguish methodological naturalism from more traditional epistemology, we picture the latter by seven theses according to Koppelberg:

(1) The methodological starting point of epistemology is the analysis of our ordinary everyday-notions about knowledge and beliefs.

(2) Epistemology makes use of terms and norms and formulates principles and aims that are not completely included in science.

(3) Epistemology has genuine philosophical methods and evidence.

(4) Epistemology has rules and norms which are logically independent from and prior to sciences.

(5) Epistemology avails not at all to scientific results and discoveries.

(6) Epistemological results (e.g. knowledge, justified beliefs) are epistemically fundamental in comparison with scientific results.

(7) Epistemology itself is logically independent and prior to science.

The more of these theses we subscribe to, the more traditional is our view about epistemology. A methodological naturalist would—at the most—opine that theses (1)-(4) are correct.

One point is—hopefully—plausible by now: It is in fact not fair to characterise all naturalists as being forced by methodological considerations to see science as the only avenue to truth (against this view, see Rea 2007, Chapter 2). Solely Quine and some admirers (see Chapter 2) hold a replacement thesis (that means the reformulation of all legitimate epistemological questions in scientific problems as an actually conducted program) that is often being criticised as unsound, overdrawn and incoherent (see e.g. Almeder 1998 or Keil 2003). The latter argues coherently that Quine—despite his rhetorical efforts—does not do empirical psychology but rather philosophy (see also Sukopp 2006a, Chapter 4).

Perhaps you are looking for a simple summarizing "formula". I cite exemplary the *Internet Encyclopedia of Philosophy*:

> 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, 1) [3]

1.2. Is naturalism a research program?

It certainly is. Perhaps it is helpful to emphasise that a research program cannot only be fruitless or "degenerative" (Lakatos), but object of revision because some of its theses turn out to be incorrect. It is no contradiction to state both: Naturalism is a research program and consists of theses (against this, see Rea 2007, Chapter 1). One example is given by Gerhard Vollmer, who advocates an ambitious naturalism: His "all-inclusive" cosmological-anthropological world-view combines post-popperian pancritical rationalism with hypothetical realism. He puts his position in twelve theses

---

[3] I would like to add two points: First, not only empirical sciences are acceptable within a naturalistic point of view. Second, naturalists disagree about the necessity of a-priori-knowledge. Philip Kitcher is a naturalist *and* thinks that we cannot disband a-priori-knowledge.

(see e.g. Vollmer 1994). The following six theses are essential. Vollmer—according to his own words—should revise his view fundamentally, if one of the following theses turns out to be deficient (Vollmer 1994, 217):

(1) Use as little metaphysics as possible.
(2) Opine for "minimal realism" according to which a world without human beings could exist. (In fact Vollmer is a hypothetical realist.)
(3) Everything consists primarily of matter, or rather, energy.
(4) Real systems are build up from simple parts.
(5) Instances which transcend all experience—that means perception plus other sources of empirical knowledge—do not exist. Therefore miracles do not exist.
(6) Cognitive efforts do not exceed nature.

The shortest possible formula of this naturalism labels it with two characteristics: *universality* (*Everywhere* in the world *everything* can be perfectly rationally explained, in German: "Alles in der Welt geht mit rechten Dingen zu") and the *restricted use of possible means*, when we try to explain or give reasons for something.

2. AGAINST QUINE: DEFENDING NATURALISM BY AFFRONTING IT

Some naturalists are carrying naturalism too far, they are too radical. I will briefly outline one famous example, Willard Van Orman Quine, which shows naturalistic incoherence or even inconsistency[4] caused by scientific

---

[4] Consistency problems are raised by calling Quine's position e.g. "self-defeating" (Moser; Yandell 1996). Accordingly, he reverts to positions that he should not accept (physicalism, empiricism). See also Almeder's (1998) incoherence thesis: after Quine's replacement thesis, all answers to legitimate epistemological questions are distilled by scientific methods when we are doing science. The point is that the replacement thesis is no completely scientific assertion, therefore self-refuting. Almeder (1998) calls it "self-defeating" and inconsistent (see also Sosa 1983; Sagal 1987 or Stroud 2001). Quine—no wonder—rejects these critiques. He argues for a kind of holism that is more complex than most critics expect. Furthermore, inconsistency seems to be a logical, not an ontological problem. Logical defects

prejudices and discrepancies between rhetoric self-ascription vs. the way Quine's own philosophy is actually carried out. In order to avoid boring the readers, I should focus on one rarely elaborated argument (2.2). First, I will have to reject Quine's famous argument for embedding epistemology in psychology (2.1). For further details cf. e.g. Hahn & Schilpp 1986, Moser & Yandell 1996, Almeder 1998, Keil 2003, Sukopp 2006a, Chapter 4.

2.1. Quine's argument for embedding epistemology in psychology

Traditional epistemology, according to the story Quine tells us, was searching for a safe fundament of human knowledge. This search was in vain. (That's by the way the shortest possible manner to simplify about 400 years of epistemological efforts. One tenet was to refute Descartes sceptical argument. All the attempts of refutations or rejections have failed. Even in mathematics knowledge is not as safe as some of us desire [see e.g. Almeder 1990, 264].) Since Hume we deal with the problem of induction. Hume shows us clearly that knowledge based on sensory data and perception also does not exist. Hume's argumentation claims that "Knowledge by Induction is not justified" (Hume 1982, 49-58), but rests on "customs" and "habits" (Hume 1982, 62), gets via Quine additional explosiveness.

If we—like Quine—neglect the distinction between analytical and synthetic propositions, and every proposition is synthetic, then "Hume's argument casts a long staggering cloud" over our efforts to give sceptics an adequate reply (Almeder 1990, 264). According to Quine, no proposition at all can claim to be safe knowledge. Therefore Quine concludes that traditional epistemology is dead. We have no "first Philosophy" and no pure philosophical truth, which justifies scientific or philosophical methods or serves as a grounding. This whole argumentation is viciously circular, because we have to assume—though we know better—a-priori-knowledge to identify and refute a-priori-knowledge.

What is really wrong with Quine's epistemology? His "settling for psychology" is a result of some deep insights (e.g. labeled by holism and

---

could not yet be demonstrated persuasively. Quine himself is not acting like a scientist. He sees place for "technology of truth-seeking" and for (restricted) norms.

underdeterminacy) and even more prejudices (behavioristic psychology, the role of social constraints in epistemology) and misinterpretations (the status of what he calls "empiricism"[5] and the role of norms). His agenda of epistemology—put in a few words—is this: If you are an epistemologist and your aim is to *clarify* the *basic grounds* of empirical sciences, than you *should* do psychology and *exert the methods of science*. (For the sake of the argument, I concede that every italicised term is clearly defined and Quine can confirm his ambitious program.) Then, Quine goes on, epistemology will be a *branch of science*, for it can only clarify—by using *scientific methods*—the *sources* and *growth* of human knowledge. Epistemology should be and will be descriptive psychology.

The most striking argument contra this "radical cure"—prescribed against the will of the "patient" philosophy—is that Quine is simply inconsistent, while he is doing what he should not do: He acts like a philosopher and only talks like a "real" naturalist. I merely mention his slogan "Epistemology should be and will be descriptive psychology". This is certainly no empirically distilled proposition, but a result of philosophical (methodological and metaphilosophical) reflection.

---

[5] Quine's notion of "empiricism" poses several questions. Here is one of them: his reference to evolution conflicts with his notion of empiricism. First, Quine is speaking of "subjective norms of similarity". Evolution, in this context, helps us to explain induction. That means to explain the assertion "Conclusions that are drawn inductively are normally correct." Quine *presupposes* the latter assumption. He does not show why induction is justified, but rather, how induction—taken as a fact—can be explained by induction. (I leave circularity-objections aside.) Furthermore the parlance of "subjective norms of similarity" contradicts Quines so called "slogan of empiricism" (Quine 1995b, 27): Nothing is in mind (has been understood) that has not been prior in the senses (i.e., accessible by sensory perception). If "subjective norms of similarity" are (parts of) innate structures of cognition, then the "slogan of empiricism" has to be revisited. If *constructive* cognitive performance of the brain is at least approximately as important as proponents of evolutionary epistemology, neuro-scientific results or philosophers like Michael Pauen, Olaf Breidbach, Gerhard Roth, Thomas Metzinger or Antonio Damasio suggest, then not too much remains of this slogan.

Finally Quine's assumptions about what science—however fallible—has found out are not true. He asserts that we *only* receive information over the world by *causal effects* of our sensory receptors (Quine 1995b, 27).

2.2. Against Quine: Relations between normativity and empiricism

I will only discuss one of Quine's falsities, namely his understanding of empiricism as the supreme norm of science. (I leave the problem of identifying "ok-sciences" aside. Quine's proceeding is quite restrictive.)

One starting point of Quine's argumentation is the following premise: Science is not accountable to any higher tribunal that may judge over science. I agree only for the sake of the argument. (In fact Quine seems to confound "science" and "reason".) Let us have a look of Quines austere view on norms. He actually sees *predictions of an observation* to test a theory as one norm (Quine 2000, 122). Further norms are conservatism, universality, simplicity, falsiability and "modesty" of a hypothesis, "anecdotes" (Quine) told by the history of science, more mathematical-formal problems, error margins and probability values (Quine 1995b, 27ff). Again, the supreme norm *is* empiricism (Quine 1995b, 29). It is part of science, and the "norm" of empiricism is also part of science and consequently fallible. That is—to put it mildly—astonishing.

The use of "norm" is not very precise. More than this: If empiricism is more sophisticated than "We do act by collecting experience, making observations or whatever good empiricists do", it is not a part of science but rather belongs to the philosophical area of understanding how scientists operate and gain their scientific knowledge. If empiricism furthermore asserts "We should act like good empiricists" (see above), then a number of norms are enclosed (e.g. "You should prefer observational data about the world compared to reasoning"). Are norms fallible? No. Some norms might be superfluous ("Men must not bear children" is a good candidate), its strict adherence sometimes has undesirable consequences ("Never lie, even if you could save a persons life if and only if you tell a lie"). A norm could be not enforceable or counterproductive. Quine's pragmatic view, combined with an absence of a keen sense of norms, forces him to underrate the relevance of norms: It is true that we need norms in the area "heuristics of obtaining hypotheses". But that is not the whole story. In a broad sense, Quine's epistemology is normative (see e.g. Beyerstein 2005). But it is insufficiently normative. What are the aims of science? How do we judge which processes are superior to other regarding a belief as knowledge? Which social factors influence science? Which *rational*

arguments should we accept on the long way from sensory radiation of our body-surface to the enormous theoretical output? These questions desire an elaborated understanding of norms.

Let us assume for one moment that these norms are dispensable. Even if we concede that Quine is initially right with most of his assumptions about empiricism, one critique remains. The data of science is the "neural input" (Quine 2000, 125). Whatever the correct way of creating theories may be, only at the edge of our "network of beliefs" our "empirical checkpoints" stay in contact with the empirically accessible world. The task of naturalistic epistemology is doing "conceptual analysis" (Quine) "within the framework of science itself" (Quine 1995a). Unfortunately, speaking of "conceptual analysis" is fairly unclear. Fortunately for philosophy, however, Quine is doing—despite his naturalistic rhetoric—philosophy, e.g. analysing meaning of terms, evaluating different philosophical theories and controversial arguments, etc. (cf. as a classical source Hahn & Schilpp 1988, or more recently Keil 2003, 253; 277ff.).

3. HOW SUCCESSFUL IS NATURALISM?—CAPABILITIES AND LIMITATIONS

Followers and opponents of naturalistic views in philosophy—be it ethics, anthropology or more theoretical philosophy—may now agree that settling for naturalism does not mean committing hara-kiri (Sagal 1987) for philosophy. Well, critics might say, that there are many ways to kill this time-honored intellectual enterprise.

To defend some naturalistic theses, I will argue that naturalism is superior to Christian philosophy (3.1), especially when we consider theism (see e.g. Nancey Murphy's "Naturalism and Theism as competing traditions" in this volume). Against Michael Rea I will try to show that even if naturalism is *only or mainly* a research program, it consists of theses, which can be identified (3.2). By "theses" I recognise propositions—be they philosophical or not—such as "The Universe only consists of matter and fields". Theses like this can be partly proven or at least criticised. Is all the work that analyses arguments against naturalism—and is based on theses at the same time—superfluous?

Finally I will sketch two limitations of naturalism: In dealing with qualia and free will, naturalism has perhaps been overextended (3.3).

3.1. Why and in what respects is naturalism superior to theism?

Nancey Murphy (see this volume) contributes a lucid—of course excursive—paper that focuses on two intellectual traditions, theism and naturalism. This is not the right place to elaborate criteria enabling us to compare long term traditions. To make allowance for historical, social or cultural effects that make these criteria "relative", e.g. "Judging the standards of argumentation in India (500 BC), you have to look at the contemporary logic (500 BC)", I concede that western ratio is not the "one and only" monolithic rational standard made for eternity. Yet theism and naturalism have to be in the same ballpark, because otherwise they cannot compete. For this reason I shall apply approved and widely accepted standards, such as explanatory power, internal and external consistency, prognostic strength, absence of circularity, etc., as straightedges for evaluation.[6]

To begin with: It is—against Murphy—*not* surprising that naturalism results in atheism or that atheism—at least agnosticism—belongs to naturalistic views. "There is no reason at all […] naturalism with regard to humans as tantamounts to atheism […]" (Murphy, in this volume, Chapter 1). There are several reasons why atheism is a consequence of naturalistic views:

   (i)   Ontological relevance of God.
   (ii)   No place for God by methodological reasons.
   (iii) The "design" of the universe (an essentially metaphysical
          argument).
   (iv)   Explanatory power of naturalism.

---

[6] I am no crusader on a mission against theism. The value of theism *for believers*, its cultural importance, a strong need of religiosity—supplying a metaphysical human need—as an anthropological constant are beyond controversy. I also concede that faith in God enables believers to have an extremely persistent, stable view of the world and even ensures social peace and harmony. My task is to distinguish between our wishes and what is—according to modern knowledge—convincing.

(i)-(iv) need of course more argumentative support than I can elaborate here. Here are some illustrative remarks on (i)-(iii).

3.1.1. Crisis of theism[7]

To start with (i): How can God show his ontological relevance? A modern ontology—be it reductive or nonreductive physicalism, emergentism, (eliminative) materialism, be it based on supervenience or not—has to decide whether it is essentially monistic, dualistic or if it grasps a plurality of entities, e.g. a realm for mental products such as "numbers" or "theoretical concepts", a realm for "concrete objects, i.e. entirely matter" or a realm of entities existing in souls and so on. If we understand "God" not as a metaphor, but rather as concrete, then his ontological status has to be explained. Murphy's access tries to stipulate theistic views with physicalism. The amazing result is consequently a version of physicalism for Christians. We are "complex physical organisms, imbued with the legacy of thousands of years of culture, and, most importantly, blown by the Breath of *God's Spirit*; we are Spirited bodies" (Murphy 2006, ix).

"Spirit" is a "bridging word", because in ordinary language we understand what we mean. "Esprit" in French e.g. means perhaps to have a clever mind. An intellectually inspiring and inspired person has "esprit". For the sake of the argument, I assume that God exists. Otherwise it is

---

[7] Unfortunately here is not the place to reply to Nancey Murphy's views on the four reasons for "Crisis in Christianity". Here is only one short hint: In chapter 2.3, "Natural Evil", Murphy tries to explain the necessity of natural evil as a by-product of natural laws using some "anthropic calculations". Well, not only Leibniz is dead. All "living-in-the-best-world-of-all-possible-world-theories" should be buried. It is true that some "[s]uffering […] is caused by their operations [natural laws; TS]." But only a cynic—or God as the most powerful cynic—would justify all natural evil as a by-product of natural laws. According to Murphy et al., this "necessary by-product of conditions [is] built into creation in order that there be creatures who could respond freely and lovingly to God […]" God could—and if he is a loving or even a most benevolent God—he should have created the world in accordance to natural laws in such a way, that almost every natural evil could be less terrifying. Is it necessary that hundred of thousands die because of floods or earthquakes? Certainly it is not. God could have created us as deterministic persons, determined to think that we have a free will. It would be much easier to respond "freely and lovingly" to God if he had "designed" a better world. "Better" of course can be identified by morally relevant terms depending on our preferred type of ethic, such as happiness or virtues.

difficult to see how he can do anything. (In fact the burden of proof lies on the shoulders of theism. Ontological possibility does not mean existence. The history of proofs for God's existence is a history of failure.) If we understand God in such a way that God can be conceived by the human cognitive system (brain), then we have to state that we have to interact with God, or vice versa. If God is essentially immaterial, then we should explain how he acts on our brains. Causality in its ordinary notion does not work here. God has to show traces of his existence, and I do not mean "miracles", but traces left after or at just the moment when we become "spirited bodies". Taken this parlance serious, we should give a proper concept for non-causal interaction or an argument that accounts for how God does act here. I would be really confident, if there was good evidence of empirical proof or—still convincing—an argument by reason.

Let us talk about (ii): A strong ontological naturalist (cf. chapter 1.1) would deny the need of any supernatural realm. Why should we accept such a view? I will just outline some arguments from Mario Bunge and Martin Mahner (2004, 226f.). A scientist (and—at the same time—a defender of theism) who does not expel supernatural entities out of the universe might say: "This experiment (or this argument) failed because the entity A has not (or has) intervened." There is no sound argument for this rather strong ad-hoc-hypothesis. Our scientist (analogous to the philosopher) might even assume that theories, predictions, etc., fail, because entity A did calibrate his measuring apparatus incorrectly. The risk of immunization against critique is obvious.

One more reason for rejecting "God" as an ontologically respectable entity is ontological austerity (see Ockham's razor): If we speak of "God" and "the devil" (e.g. as an fallen angel), then we could also accept the devil's grandmother and other entities leading to ontological rank growth (see Bunge & Mahner 2004, 11).[8]

Furthermore we and the observable parts of our world are structured. We observe regularities and are able to predict the behavior of natural systems. If we assume supernatural entities, then by definition we could not have any access to them. Apart from miracles, we could not know

---

[8] Serious ontological problems of e.g. naturalism as materialism remain: How does Beethovens 9th symphony exist? How do numbers exist?

anything about such entities, because they are not bound by mundane regularities, e.g. laws of nature.

Last but not least, supernatural Gods, angels or demons could explain everything. A (scientific) theory should have great explanatory power, i.e., it should explain *a couple of* things but not everything. A theory that explains everything also does not exclude anything and consequently has no explanatory power (less excursive than the last two passages are Mahner 2002; Bunge & Mahner 2004).

I will now outline argument (iii): The assumption of a "sparingly furnished universe" as a consequence of an naturalistic ontology allows some wide implications. Like Franz Josef Wetz, we can see fundamental questions about our world-view ("weltanschauliche Fragen") following a "more metaphysic naturalism" (Wetz 2003a, 42). Human beings are—in this view—unimportant agents in an aimless and senseless universe that is ruled by blind natural forces. This senseless universe (nature) is everything that exists. "It is the look, the design[9] or the physiognomy of the physically constituted universe" that make the old notion called creation loosing its plausibility (Wetz 2003b, 70). Though the existence of God is logically possible, there is no sound argument for this pure possibility. We see some serious consequences interpreting ontological aspects (the "furniture of the universe", or of the world we are living in; see e.g. Wetz 1994; Kanitscheider 1995). To put it in the words of Kanitscheider (1995, 67):

> Today there is strong evidence for the estimation, that all these pluralistic ontologies, which had the historical function, to separate men in at least some sub-functions from the context of nature, were *metaphysical illusions* [italicised in the original; TS], corresponding to an urgent desire of sense. Under the pressure of [scientific facts these illusions cannot be maintained].

3.2. Naturalism: A yielding, progressive research program?

My task is "only" to show that a kind of methodological naturalism is in fact a progressive research program. My standards for success are: Explanatory power, capacity of problem solving and potential to integrate

---

[9] "Design" in this notion should be understood in a broader sense, and not the way it is found in current fields of discussion like "intelligent design", creationism or teleological theories in biology and so on.

new questions and fields of research. One question always is: Given these criteria, is a naturalistic approach more successful than a competitive one?

Again, methodological naturalism in this context means that scientific (and other!) *methods* and *results* are indispensable for philosophers. Against Michael Rea (Chapter 3 of his paper in this volume), naturalism is not "inextricably tied to scientific realism by virtue of treating the methods of science as basic sources of evidence" (against this view see e.g. Sukopp 2006a). Naturalists are sometimes instrumentalists or prefer an internal realism or a hypothetical realism. A naturalist like Gerhard Roth prefers constructivism—which is, in the philosophical landscape, not too far from realism!

Second, the ontological austereness of naturalism does not force it to neglect everything beyond physical objects (see e.g. Bunge & Mahner 2004). Science does not solve all the ontological problems (see e.g. Searle 1997).

My argument in favor of naturalism shows that naturalists are not tied to "scientific realism", but have to consider science as relevant for philosophy. Ulrich Frey (see his paper in this volume) gives an example for a naturalistic Philosophy of Science that is based on empirically accessible data on cognitive abilities of scientists. Contrary to Ulrich Frey, I have serious doubts that such a Philosophy of Science (labeled the "Naturalistic account of scientific errors and its consequences"; see Chapter 3 of Frey's paper) tells us the whole story. But in most respects the cognitive approach is fruitful, because of

  (i)   its explanatory power;
  (ii) its power of problem solving and
  (iii) its capacity of integrating new question into an old discipline
        of philosophy.

(i): A cognitive approach chosen by Frey could show empirically why scientists underlie mechanisms such as framing effects or failure by deficient reducing of complexity. What we know through cognitive science (cognitive psychology) is relevant in Philosophy of Science. Top-down, all-over rules for instance, like the methodology of critical rationalism, should consider cognitive abilities. We do have problems leaving the

paradigm of well-tried theory behind and therefore hesitate and refuse to give new theories a chance.

(ii): Historical case studies give strong evidence in support of the thesis that "natural science like cognitive psychology and evolutionary biology can offer good *descriptions* and *explanations* of phenomena that are of interest in the philosophy of science" (Frey 2007, beginning of Chapter 4; italicised in the original, TS). This cognitive approach is a more powerful instrument than e.g. rational reconstruction or other, rather formal, criteria to evaluate and compare theories. It shows that scientists are subject to cognitive restrictions, it analyses what types of limitations are relevant and it offers evolutionary explanations. Evolution should not be "the instrument par excellence", because if your only instrument is a hammer, then the whole world looks like a nail. To use another metaphor: You hit the wall with your hammer, and after damaging it heavily, you have finally banged the nail into the wall. Really convinced, you declare: "Look at my proper, deliberately chosen tool". Let us stick to facts again. If brains are evolved systems underlying the same processes—like selection and adaptation—, if we take evolutionary theories to be the best theories that explain biological limits, and if cognition is entirely biologically determined, then the conclusion is unspectacular: failures, (cognitive) prejudices, beaten paths exert influence on scientific practice.

Now to (iii): It is widely accepted that certain disciplines—like biology or psychology—are philosophically relevant, because they raise similar questions. One of these questions is: By which processes do we acquire (justified) beliefs? I would like to underline the integrating potential of a cognitive approach quoting Hilary Kornblith (1994):

> Even granting for the sake of the argument that in principle it is possible to answer epistemological questions a priori, epistemological truths are anything but obvious. It would be foolhardy not to subject epistemological theories to empirical tests. If scepticism is to be rejected, then epistemology and psychology impose significant constraints on each other. The best way to develop epistemological theories is thus to employ these constraints in a way that allows us to prod the theory along by confronting it with empirical tests.

Moreover: psychology and epistemology could mutually benefit from an approach that makes allowance for empirically accessible factors like a)

invariance of errors or b) universality of some cognitive mechanisms such as linear-problem-solving, neglecting side effects, long term effects or feedback-loops (Frey 2006, 89; Frey 2007, Chapter 5). The agenda of epistemology contains at least three questions:

(1) By which processes do we acquire (justified) beliefs? (see above).
(2) By which processes should we acquire (justified) beliefs?
(3) Are the processes by which we acquire (justified) beliefs the same processes by which we should acquire (justified) beliefs?

No doubt the answers to these questions depend on the (degree of) reliability of cognitive processes.

3.3. Limitations of naturalism

Martially phrased, a naturalist should show which territories he is actually able to control with his troops.[10] It is out of place to think that naturalism is

---

[10] In Sukopp 2006b I offered an overview by classifying antinaturalistic arguments. I suggested that the best arguments attack fundamental naturalistic premises (namely methodological, epistemological and metaphilosophical arguments). They criticise (scientific) realism or some normative premises that cannot be justified within science itself. Finally, we find on the agenda of some critics the status of reason. We are in the midst of old disputes. Those critics who disagree with naturalistic premises pick up the following problems (see e.g. Putnam 1982; Keil 1993; most of the papers in Bartelborth 1996; Moser & Yandell 1996; Haaparanta 1999; Hartmann & Lange 2000; Goebel 2003 (2005); Graefrath 2005; Loeffler 2005; Wagner & Warner 2005):
    1. Inconsistency.
    2. Circularity of conclusions.
    3. Missing or at least inadequate opportunity to consider normativity.
    4. Neglect of any form of apriorism (a moderate apriorism is hold to be true by some naturalists, e.g. Kitcher).
    5. Antipsychologism.
    6. (Strong) replacement thesis (attacked also by naturalists, e.g. Almeder).
  I offer the following scheme of problems resulting from naturalistic premises:
    1. Against realism (not all naturalists prefer a certain version of realists): constructivism, instrumentalism, idealism.

a solely successful affair. In this paper I only can outline two examples that show naturalistic failures or—which would be a more striking argument against naturalism—intrinsic limitations.

3.3.1. The naturalisation of recalcitrant phenomena: Qualia

Qualia behave like a stubborn horse in the eye of a naturalist. (Some naturalists behave like Cowboys trying to ride a stubborn horse.) What it is like in first-person-view to feel pain or to play ping-pong is not the same as to look at fNMR-diagrams (functional Nuclear Magnetic Resonance), a modern and powerful instrument to look "inside the brain". What does this mean? It means that we see "neuronal correlates" to more or less active areas of the brain, which are—e.g. in the case of fNMR—oxygenated. What it means to say "I feel pain" is part of ordinary experience. Science may help to understand what cerebral processes are necessary to say "I feel this tooth aching", but cerebral processes alone do not pinpoint sufficient conditions for this highly theoretical statement. We cannot locate the area that is "responsible" for saying "I" or for knowing if we are justified to say "I". Perhaps Thomas Metzinger (1995, 2003, 2004) is right by pointing out that there is no instance or centre "self" that says "I", but rather a system of self-representations of a certain type. I do not think that naturalism is fruitless in this area. I am simply of the opinion that naturalists should keep in mind what philosophical conclusions we really can draw from empirical results. That means, for instance, that phenomenological and other non-naturalistic views cannot be denied.

3.3.2. A remark on neurology and philosophy of free will

Shouting "Eureka" sometimes makes the shouter feel better. The sweeping arising from the feature pages of well-known German newspapers occasionally sounds like "Eureka". Free will was even worth making the headlines. Why? Because of a bold assumption that "the" problem of free

---

2. Against terminological pre-decisions, e.g. the naturalistic notion of nature (see e.g. Keil 1993, 360: "Who does not like in no case to speak about 'nature' should remain silent about naturalism.").

3. Against the status of reason/experience: empiricism-rationalism debates, externalism vs. internalism.

will is empirically solved (A). No such free will exists. Again, I do not neglect efforts made in the "Neuro-Philosophy of free will" only because I reject thesis A.

There is no doubt that people can be "influenced to act in a certain way". But that is not the same as to demonstrate if an action is performed according to someone's free will (choice) or if it is not.

I really do not know of any experiment that gives strong evidence of argumentative leave-taking of free will. I have to insist—though I am not convinced that we have some capability to "decide according to our own free will"—that the adequacy of free will-concepts has not yet been empirically decided. My task is just to sketch a sufficiently complex situation for evaluating if person A acts according to his or her "free will". This situation should not be simplified arbitrarily just because the quantity of potential parameters in psychological experiments *has to be restricted* or the situation *can only be accessible* to neural-psychological interpretation if it is simplified. More precisely: An experiment that gives evidence for the strong thesis "Free will is just an illusion" should meet the following requirements:

(1) Against his/her (rational) preferences, motivations, etc., person A chooses to perform an action, e.g. action 1.

(2) Theoretically A could have decided to take alternative 2. We need two (or more) alternatives (1, 2, 3, …) which are—at first sight—on a par.

(3) Neural correlates of the relevant brain areas—such as the limbic system—should be identified to demonstrate the ineffectiveness of preferences (motivation) mentioned in 1. (Not to have free will means e.g. that these more or less rational preferences do not take effect.)

(4) Before the determined action (the action that shows the absence of "free will") is carried out, I would like to see neural correlates of deliberations such as preferences, motivation (see again 1 in this list). After this, a determined—e.g. preconscious process—assumes the leadership.

(5) What does "action potential" mean in this context? Can we properly correlate "action potential" to "carrying out the action"? I invite scientist to think more about the meaning of terms ("action", "influence", "action potential"), because they are suggestive.

Denying free will *and* remaining silent about limiting experimental preferences, deliberations about which alternative we could have chosen in the experiment (see again 1 in the list), is suspicious. Such an approach is not philosophically reflected, because it anticipates interpretations.

One remark concerning interpretations: An experiment that suffices conditions (1) to (5) has to be interpreted with semantic caution. Winfried Löffler[11] gave an impressing example of how slight (and obvious) misunderstandings, wrong quotations, dubious reinterpretations of terms and so can on lead to "Eureka": "We can demonstrate empirically that free will is an illusion."

To put it in one sentence: Current experiments are much too simple to simulate real-life situations in which actions could be carried out through someone's free will (or not).

4. CONCLUSION

I thus summarize:

(1) Naturalism covers a wide range of positions, which can be identified as theses within research programs or as claims like "We can interact with all existing entities". This claim means the rejection of a supernatural sphere or realm. A naturalist sees cognitive processes, and therefore epistemological results, as natural phenomena, partially explainable by sciences (and humanities!). That is one short formula for naturalism on the different levels: ontology, methodology, epistemology (in the narrower sense).

---

[11] I refer to Löffler's talk at the 29th Wittgenstein-Symposium on 12th of August 2006, titled "What naturalists always knew about freedom: A case study in the narrative sources of 'scientific facts'".

(2) Naturalism does not rely on trivialising philosophy, neither by saying "Science itself teaches us", nor by misconceiving naturalism as a self-evident position such as "Philosophers should sometimes take scientists seriously".

(3) Only a few naturalists—and unfortunately some more non-naturalists—still think that Quine is a standard naturalist or even its prototype. Quine's naturalised epistemology is problematic in many respects. He is incoherent, or even worse, inconsistent, cannot face normative challenges, misunderstands "empiricism" as a norm, etc.

(4) I argue that we can disband theism in comparison with naturalism. Among a lot of "rearguard action", one move of contemporary theism is the assertion that tries to reconcile theism with naturalism. Theism and naturalism dissent. They cannot both be true! In respect of explanatory power, economy, and plausibility, as well as for several other reasons, naturalism is superior to theism.

(5) Where is naturalism successful? Cognitive approaches to Philosophy of Science and epistemology give examples for philosophical relevance of scientific methods and results.

(6) Philosophers are no scientists and some scientific shortcomings are misleading. Anyone who claims that "free will" is dispensable due to experiments that prove this should pause for a moment and think e.g. about the adequacy of the experimental situation.

(7) To answer the question of this volume: Yes, naturalism is successful in some respects. Yet if we adopt some ultimate criteria for success (see Rea 2007, Chapter 1), like "immediate acquisition of wealth and happiness", simply because of "adopting" naturalism, then naturalism is not successful. By the way, it is comforting for naturalists—and particularly for their opponents—that "non-naturalists [do not] spontaneously burst into flames" (Rea 2007, Chapter 1).

5. REFERENCES

Almeder, R. 1990: "On naturalizing epistemology." *American Philosophical Quarterly* 27, 263-279.

Almeder, R. 1998: *Harmless Naturalism: the limits of Science and the Nature of philosophy*. Chicago, La Salle: Open Court.

Audi, R. 1998: *Epistemology: a contemporary introduction to the theory of knowledge*. London, New York: Routledge.

Bartelborth, T. 1996: *Begründungsstrategien: Ein Weg durch die analytische Erkenntnistheorie*. Berlin: Akademie Verlag.

Beyerstein, L. 2005: *Why Naturalized Epistemology is Normative*. http://majikthise.typepad.com/majikthise_/files/en_20050108_blog.pdf (17.05.2007).

Bunge, M. and Mahner, M. 2004: *Über die Natur der Dinge. Materialismus und Wissenschaft*. Stuttgart: Hirzel.

Frey, U. 2006: "Is naturalism progressive? A naturalistic approach to the Philosophy of Science." In: Gasser, G. and Kanzian C. (eds.) *Papers of the 29th International Wittgenstein Symposium "Cultures: Conflict-Analysis-Dialogues."* Kirchberg am Wechsel: Austrian Ludwig Wittgenstein Society, 88-90.

Frey, U. 2007: "Naturalized Philosophy of Science: A cognitive approach." In: Gasser, G. (ed.) *How successful is naturalism?* Frankfurt am Main: Ontos.

Goebel, B. 2003: "Probleme eines philosophischen Naturalismus." In: Goebel, B., Hauk, A. M. and Kruip, G. (eds.) 2005, 23-42.

Goebel, B., Hauk, A.M. and Kruip, G. (eds.) 2005: *Probleme des Naturalismus: Philosophische Beiträge*. Mentis: Paderborn.

Goldman, A. 1994a: "What is Justified Belief?" In: Kornblith, H. (ed.) *Naturalizing Epistemology*. Cambridge, MA: MIT Press, 105-130.

Goldman, A. 1994b: "Epistemic Folkways and Scientific Epistemology." In: Kornblith, H. (ed.) *Naturalizing Epistemology*. Cambridge, MA: MIT Press, 291-316.

Goldman, A. I. 2006: "Social Epistemology." *Stanford Encyclopedia of Philosophy* (substantively revised version). http://plato.stanford.edu/entries/epistemology-social/ (17.05.2007).

Gräfrath, B. 2005: "Biologismus in der Ethik und Erkenntnistheorie." In: Goebel, B., Hauk, A.M., and Kruip, G. (eds.) *Probleme des Naturalismus: Philosophische Beiträge*. Paderborn: Mentis, 101-110.

Haaparanta, L. 1999: "On the possibilities of naturalistic and of pure epistemology." *Synthese* 118, 31-47.

Hahn, L. E. and Schilpp, P. A. (eds.) 1988: *The Philosophy of W.V. Quine*. La Salle, IL: Open Court (The Library of living Philosophers Vol. XVIII).

Hartmann, D. and Lange, R. 2000: "Ist der erkenntnistheoretische Naturalismus gescheitert?" In: Keil, G. and Schnädelbach, H. (eds.), 144-162.

Hedrich, R. 1998: *Erkenntnis und Gehirn: Realität und phänomenale Welten innerhalb einer naturalistisch-synthetischen Erkenntnistheorie*. Paderborn: Schöningh.

Isak, R. (ed.) 2003: *Kosmische Bescheidenheit. Was Theologen und Naturalisten voneinander lernen können*. Freiburg: Katholische Erzdiözese Freiburg.

Kanitscheider, B. 1995: *Auf der Suche nach dem Sinn*. Frankfurt am Main and Leipzig: Insel.

Keil, G. 1993: *Kritik des Naturalismus*. Berlin and New York: De Gruyter.

Keil, G. and Schädelbach, H. 2000: "Naturalismus." In: Keil. G. and Schnädelbach, H. (eds.) *Naturalismus: Philosophische Beiträge*. Frankfurt am Main: Suhrkamp, 7-45.

Keil, G. 2003: "'Science itself teaches'. A fresh look at Quine's naturalistic metaphilosophy." *Grazer philosophische Studien* 66, 253-280.

Koppelberg, D. 2000: "Was ist der Naturalismus in der gegenwärtigen Philosophie?" In: Keil, G. and Schnädelbach, H. (eds.), *Naturalismus. Philosophische Beiträge*. Frankfurt am Main: Suhrkamp, 68-91.

Kornblith, H. 1994: "Introduction: What is Naturalistic Epistemology?" In: Kornblith, H. (ed.) *Naturalizing epistemology*. Cambridge, London: MIT Press, 1-11.

Kornblith, H. 2002: *Knowledge and Its Place in Nature*. Oxford: Clarendon Press.

Koskinen, H. J. 2004: *From a Metaphilosophical point of view. A study of W. V. Quine's Naturalism*. (Acta Philosophica Fennica Vol. 74). Helsinki: Hakapaino Oy.

Laudan, L. 1987: "Progress or Rationality? The Prospects for Normative Naturalism." *American Philosophical Quarterly* 24, 19-31.

Laudan, L. 1996: *Beyond Positivism and Relativism : theory, method and evidence*. Boulder, CO: Westview.

Löffler, W. 1999: "Naturalisierungsprogramme und ihre methodologischen Grenzen." In: Quitterer, J. and Runggaldier, E. (eds.) *Der neue Naturalismus – Herausforderungen an das christliche Menschenbild*. Stuttgart: Kohlhammer, 30-76.

Löffler, W. 2005: "Evolutionäre Ethik und Erkenntnistheorie: Methodologische Voraussetzungen." In: Goebel, B., Hauk, A. M. and Kruip, G. (eds.) *Probleme des Naturalismus: Philosophische Beiträge*. Paderborn: Mentis, 111-125.

Mahner, M. 2002: "Stichwort: Naturalismus." *Naturwissenschaftliche Rundschau* 155, 689-690.

Metzinger, T. 2000 (ed.): *Neural Correlates of Consciousness — Empirical and Conceptual Questions*. Cambridge, MA: MIT Press.

Metzinger, T. 2003: *Being No One. The Self-Model Theory of* Subjectivity. Cambridge, MA: MIT Press.

Metzinger, T. [3]2004: *Subjekt und Selbstmodell. Die Perspektivität phänomenalen Bewußtseins vor dem Hintergrund einer naturalistischen Theorie mentaler Repräsentation*. Paderborn: Mentis.

Moser, P. K and Yandell, D. 1996: "Against Naturalizing Rationality." *ProtoSociology (special issue on rationality)* 8, 81-96.

Murphy, N. 2006: *Bodies and Souls, or Spirited Bodies?* Cambridge: Cambridge University Press.

Murphy, N. 2007: "Naturalism and Theism as Competing Traditions." In: Gasser, G. (ed.) *How successful is Naturalism?* Frankfurt am Main: Ontos.

Putnam H. 1982: "Why reason can't be naturalized." *Synthese* 52(1), 3-23.

Quine, W. V. O. 1995a: *From Stimulus to science*. Cambridge, MA: Harvard University Press.

Quine, W. V. O. 1995b: *Unterwegs zur Wahrheit: konzise Einleitung in die theoretische Philosophie*. Paderborn: Schöningh.

Quine, W. V. O. 2000: "Naturalismus—oder Nicht über seine Verhältnisse leben." In: Keil, G. and Schnädelbach, H. (eds.) 2000, 113-127.

Rea, M. C. 2005: "Naturalism and Ontology: A Reply to Dale Jacquette." *Faith and Philosophy* 22, 343-357.

Rea, M. C. 2007: "How Successful Is Naturalism?" In: Gasser, G. (ed.) *How successful is Naturalism?* Frankfurt am Main: Ontos.

Sagal, P. T. 1987: "Naturalistic Epistemology and the Harakiri of Philosophy." In: Shimony, A. and Nails, D. (ed.) *Naturalistic epistemology*. Dordrecht: Reidel (Boston studies in the Philosophy of science; v. 100), 321-332.

Searle, J. R. 1997: *Die Konstruktion der gesellschaftlichen Wirklichkeit. Zur Ontologie sozialer Tatsachen*. Reinbek bei Hamburg: Rowohlt.

Sosa, E. 1983: "Nature unmirrored, Epistemology naturalized." *Synthese* 55, 49-72.

Stroud, B. 2001: "The significance of Naturalized Epistemology." In: Dagfinn, F. (ed.) *Philosophy of Quine*. London: Garland, 455-471.

Sukopp, T. 2006a: *Naturalismus — Kritik und Verteidigung erkenntnistheoretischer Positionen*. Frankfurt am Main: Ontos.

Sukopp, T. 2006b: "Outline of Arguments against Naturalism: Levels and Classifications." In: Gasser, G. and Kanzian C. (eds.) *Papers of the 29th International Wittgenstein Symposium "Cultures: Conflict-Analysis-Dialogues."* Kirchberg am Wechsel: Austrian Ludwig Wittgenstein Society, 344-347.

Vollmer, G. 1994: "Was ist Naturalismus? Zwölf Thesen zur Begriffsverschärfung." *Logos, N.F.* 1, 200-219.

Wagner, S. and Warner, R. 2005: "Naturalismus — Eine kritische Würdigung." In: Goebel, B., Hauk, A.M. and Kruip, G. (eds.) 2005, 43-61.

Wetz, F. J. 1994: *Lebenswelt und Weltall. Hermeneutik der unabweislichen Fragen*. Stuttgart: Neske.

Wetz, F. J. 2003a: "Der neue Naturalismus. Eine Annäherung." In: Isak, R. (ed.) 2003, 41-66.

Wetz, F. J. 2003b: "Naturalistische Bescheidenheit? Kritische Anfragen an das theologische Menschenbild." In: Isak, R. (ed.) 2003, 67-82.

Wrenn, C. B. 2006: "Naturalistic Epistemology." *The Internet Encyclopedia of Philosophy*.
http://www.iep.utm.edu/n/nat-epis.htm (23.10.2006).

Part II

Epistemology and Ontology

How Successful Is Naturalism?

Michael C. Rea
University of Notre Dame

The question raised by this volume is "How successful is naturalism?" The question presupposes that we already know what naturalism is and what counts as success. But, as anyone familiar with the literature on naturalism knows, both suppositions are suspect. To answer the question, then, we must first say what we mean in this context by both 'naturalism' and 'success'. I'll start with 'success'. I will then argue that, by the standard of measurement that I shall identify here, naturalism is an utter failure.

1. STANDARDS FOR SUCCESS

So what would it be for naturalism to be *successful*? Well, it could be any of a number of things. Naturalism would surely enjoy a kind of success if everyone were to become a naturalist, if becoming a naturalist were regularly and reliably attended by the immediate acquisition of wealth and happiness, if all and only non-naturalists were to spontaneously burst into flames, and so on. But whatever value might attach to these kinds of success, I doubt that they are what most of us are interested in when we inquire after the success of naturalism. Our aims are more modest. When we ask about the success of naturalism, I suspect that what we primarily want to know is whether it is subject to serious objections. If it is not, then it is at least modestly successful. If it is not and its rivals are, all the better.

Two related objections are commonly taken to devastate naturalism: first, that it is self-refuting; second, that it has rationally unacceptable consequences. For reasons that shall become clearer below, I think that both of these objections are non-starters—far from devastating, they aren't even in the right ballpark. So I'll leave them aside. What I want to focus on is a third sort of problem: the malady that a philosophical position suffers from when it rationally commits its adherents to views that are in direct

tension with whatever attitudes, goals, or values partly characterize the position. For lack of a better label, call this malady *dissonance*.

Not every philosophical position is characterized in part or whole by values, attitudes, or goals. Lewisian modal realism is a philosophical position, but there is no reason at all to think that anything like goals, values, or attitudes are even partly constitutive of that position. One reason for this is that there is no tradition associated with the position. The set of Lewisian modal realists almost certainly has fewer than ten members; and what unites them is arguably nothing more than adherence to the central thesis of *On the Plurality of Worlds*. Positions like empiricism, materialism, rationalism, and the like are different, however. Even if there are specific philosophical theses that express these positions, part of what it is to *be* an empiricist or a materialist or a rationalist is to manifest certain attitudes, goals, or values. As Bas van Fraassen (2002, 58ff.) notes, adherents of these positions "know how to retrench" when developments in science or philosophy come into tension with theses they endorse, and it is the attitudes, goals, and values that serve as guides to the retrenchment.

It is not irrational to embrace a dissonant position, but it is unpleasant. Moreover, the role played by unpleasantness in our decisions about which philosophical positions to adopt is not to be underestimated. Dissonant theories fail along the dimensions of elegance and conservatism: by definition, they demand revision of pre-philosophical attitudes and values, and they break with traditions that we take seriously as guides to the positions we adopt. It is widely acknowledged that elegance and conservatism, among other pragmatic virtues, play an important role in scientific theory choice, and there is no reason to doubt that they play similar roles in philosophical decision making. Indeed, there is every reason to think that their role is all the more vital in decisions about philosophical positions that cannot be formulated as theses—positions like empiricism and materialism as they are understood by van Fraassen (of which more below), and like naturalism as I shall characterize it.

Thus, to accuse a position of being dissonant is to level a serious charge against it. The charge isn't as serious as self-defeat, or incoherence. Nevertheless, a position that is dissonant fails along an important dimension for measuring success. It is in this way, so I shall argue, that naturalism is a failure.

My argument has the form of a dilemma: If naturalism is a thesis, it is dissonant. If it is not a thesis, it avoids one kind of dissonance but falls prey to another. Either way, then, it is dissonant.

2. THE NATURE OF NATURALISM

In my book *World Without Design*, I argued that naturalism is not a philosophical thesis, but a research program. A research program is a set of methodological dispositions—dispositions to treat certain kinds of arguments or belief-sources as basic sources of evidence. On my view, naturalism is a *shared* research program—a subset of a maximal set of methodological dispositions[1]—that treats the methods of science and those methods alone as basic sources of evidence. Among its most important rivals are intuitionism and supernaturalism, which differ only by treating certain *additional* sources as basic (intuition in the case of intuitionism and religious experience in the case of supernaturalism). A source of evidence is treated as basic just in case it is trusted without reliance on independent evidence in favor of its reliability.

The claim that naturalism is a research program is controversial. But, then again, so is any view about the nature of naturalism: there is really no consensus about what exactly the position involves. Some say that naturalism is a metaphysical view (for example: the view that the universe is a closed causal system). Others say that it is an epistemological view (for example: the view that scientific inquiry is the only avenue to knowledge). Still others say that it is a view about philosophical methodology (for example: the view that philosophers ought to abandon traditional problems about skepticism and ontology and pursue their various projects in a way continuous with the methods of science.) There is broad consensus that, whatever it is, naturalism involves high regard for the methods of science and low-regard for non-scientific modes of theorizing. But a precise and even modestly non-contentious statement of what more is involved in naturalism has yet to appear.

---

[1] A set of methodological dispositions is maximal just in case it is possible to have all of the dispositions in the set but it is not possible to have all of them and to have other methodological dispositions as well.

Some say that naturalism comes in different *varieties*, each expressible by a different philosophical thesis. The typical varieties listed are metaphysical, epistemological, and methodological naturalism. My own view, however, is that there is in fact only *one* version of naturalism, and many *mischaracterizations* of it. Given the current state of the literature, to say this is to say that many naturalists have mischaracterized their own naturalism. But I think that there are very good reasons for making this apparently uncharitable claim. In particular, I think that it is the only way for naturalism to avoid a certain kind of dissonance.

Naturalists are united at least in part by dispositions that preclude allegiance to views that cannot be called into question by developments in science. Part of what it is to be a naturalist is to respect the methods of science above all other forms of inquiry and to manifest a disposition to follow science wherever it leads. But if we take this idea seriously, then we are led fairly directly to the conclusion that naturalism *couldn't* be a substantive philosophical thesis. It is clear that, if naturalism were a thesis, it would be a thesis of metaphysics, epistemology, or philosophical methodology. But the consensus among naturalists is that, in matters of metaphysics, epistemology, and methodology, all of our theories must ultimately be justified by the methods of science, any of our theories might be overthrown by science, *and* one must follow science wherever it leads. So, on the one hand naturalists are committed to following science wherever it leads; on the other hand, they are committed to thinking that following science wherever it leads might force one to reject any thesis that might plausibly be identified with naturalism. To regard naturalism as a *thesis*, then, is to suppose that what is really central to naturalism is dogmatic adherence to some view in metaphysics, epistemology, or methodology—such that if the view in question were overthrown by science, naturalists would not retrench, rather they would be refuted. But if that supposition is correct, naturalism is dissonant. For dogmatic adherence to any thesis of metaphysics, epistemology, or methodology is in direct tension with the sort of respect for science and disposition to follow science wherever it leads that lies at the heart of the naturalist tradition.[2]

---

[2] The argument in this paragraph is a slightly modified version of an argument I gave in *World Without Design,* 51ff. The modifications were inspired by an argument

To avoid dissonance, then, naturalism must be characterized as something other than a thesis. I suppose there are many other things that it *could* be: an attitude, a value, a preference, a stance, etc. But it is not clear to me that there is much difference between saying that naturalism is one of these things and saying that it is what I say it is—namely, a research program. At any rate, what does seem clear—and what is most important for present purposes—is that naturalism is best characterized as something *other* than a thesis; and whatever label one wants to apply to it, what it seems to be most centrally is a plan or disposition to use the methods of science and those methods alone in the development of philosophical theories.

As I have said, this view of naturalism is controversial. But others have held it.[3] Moreover, as I argued at length in *World Without Design*, taking naturalism this way fits very nicely with characterizations offered by the most prominent spokesmen for the naturalist tradition in the 20th Century, John Dewey and W. V. Quine. And, furthermore, this view of naturalism faithfully captures what is common to virtually all who call themselves naturalists without immediately rendering naturalism vulnerable to the charge of dissonance.

3. DISSONANCE FROM ANOTHER SOURCE

In characterizing naturalism as a research program, I have rendered it immune to a variety of objections. It is not a thesis, so it is not refutable. (And so, for this reason, it cannot be self-refuting as is commonly alleged.) Research programs can have consequences—the consequences of a research program are just those theses to which one is rationally committed by virtue of adopting the research program fully, consistently, and competently. But research programs do not, strictly speaking, imply anything. So one can't refute naturalism by showing that it entails a falsehood. One might hope to show that naturalism has consequences that are rationally unacceptable, but the hope is in vain. For any such maneuver

---

developed independently by Bas van Fraassen for the conclusion that empiricism is not a philosophical thesis but a "stance". See van Fraassen 2002, Ch. 2, esp. 35 – 46.

[3] To take just two examples, see Sellars 1922, vii, and Forrest 1996, 89.

is destined to be dialectically ineffective. Again, the consequences of a research program are just those views to which one is *rationally* committed by virtue of adopting it. So it will be futile to try to convince a naturalist that she should regard the consequences of naturalism—views to which she is in fact rationally committed—as rationally unacceptable. Nevertheless, research programs can still prove dissonant. And this is what we find in the case of naturalism.

Preliminary to showing this, two further features of naturalism must be brought to light. First, unlike empiricism, the close cousin with which naturalism is often mistakenly identified, naturalism is inextricably tied to scientific realism by virtue of treating the methods of science as basic sources of evidence. Precisely because they regard the methods of science as *evidential* sources, naturalists are committed to thinking that those methods are reliably aimed at truth and that the theories produced by those methods are worthy of belief. This is scientific realism—or, at any rate, it is one variety thereof.

Second, naturalists almost universally take themselves to be committed to an ontology that includes only things that can be investigated by science. This is most evident in the various slogans that have been offered flippantly or in earnest as characterizations of naturalism or of what is sometimes called 'metaphysical' or 'ontological' naturalism. Quine, for example, characterizes naturalism as 'the recognition that it is within science itself, and not in some prior philosophy, that reality is to be identified and described' (1981, 66), the implication being that the correct ontology just *is* the ontology of science. Likewise, Frederick Schmitt, in the *Blackwell Companion to Metaphysics* entry on naturalism, characterizes *ontological* naturalism as the view that only natural objects are real, where 'natural' is understood to refer to whatever is recognized by science. (1995, 343) Wilfrid Sellars has famously said that "Science is the measure of all things, of what is that it is, and of what is not that it is not." And Armstrong's characterization of naturalism as the view that reality consists of "nothing but a single, all-embracing spatiotemporal system" (1980, 35) seems clearly motivated by commitment to an ontology including nothing beyond objects that can be investigated by science. Of course, these are but a few examples. But as anyone familiar with the literature on naturalism will attest, they are perfectly representative.

These two commitments lay at the heart of the naturalist tradition. But the trouble is that there is direct tension between these two commitments on the one hand and some of the consequences of naturalism on the other. It is to a defense of this claim that I now turn.

In *The Empirical Stance* (and elsewhere), Bas van Fraassen distinguishes empiricists from metaphysicians as follows: Metaphysicians give absolute primacy to demands for explanation, and are satisfied with explanations by postulate; empiricists reject demands for explanation at certain points, and are unhappy with explanations by postulate. (2002, 36ff.) In light of van Fraassen's other work, I am inclined to gloss this distinction as follows: The metaphysicians are those for whom explanatory power is an *important* theoretical virtue, and an *epistemic* virtue. For metaphysicians, inference to the best explanation (whatever exactly that comes to) is a reliable way of acquiring true beliefs about the world, and if a phenomenon is best explained by the postulation of *x*s, then one ought to believe in *x*s. For empiricists, matters are otherwise: explanatory power is not terribly important, it is not an epistemic virtue, and the fact that postulating *x*s best explains some phenomenon is not much of a reason (if it is any reason at all) for believing in *x*s. Now, a question: Does naturalism take its stand with the metaphysicians, or with the empiricists?

We might appeal to authorities (like Dewey, Quine, Armstrong, and others) to try to settle this question. But, really, we don't have to since naturalism's commitment to scientific realism settles the question for us. If naturalism were to take its stand with the (van Fraassen-style) empiricists, it would be committed to thinking that the methods of science are *not* reliably aimed at truth. Why? Because inference to the best explanation plays an important role in scientific theorizing, and appeal to explanatory power plays an important role in scientific theory choice. Thus, to *deny* that explanatory power is an epistemic virtue is precisely to affirm that scientific theories are sometimes (maybe often) chosen for reasons that are not correlated with likelihood of truth. And in that case, it makes no sense to believe a theory *simply* because it has been selected by scientific methods. So naturalists, by virtue of their commitment to scientific realism, must take their stand with the metaphysicians: Explanatory power *is* an epistemic virtue; inference to the best explanation provides reason for belief.

But now comes trouble. In *World Without Design*, I argued that naturalists are committed to, among other things, some form of substance dualism. The argument for this conclusion can be summed up as follows:

(i)   Naturalists are committed to believing only what can be justified via the methods of science. But,

(ii) though the methods of science provide justification for believing in material objects and for believing that material objects have modal properties, either (a) those methods provide no justification for believing that the modal properties of material objects are intrinsic, or (b) they do so only via pragmatic arguments.

(iii) Constructivism—the thesis that modal properties are mind-dependent—provides the best explanation for our modal knowledge if (iia) is true; and it also provides the best explanation for the truth-conduciveness of pragmatic arguments. Thus

(iv) if (iia) is true, then naturalists should accept constructivism; and if (iib) is true, naturalists should accept constructivism. But

(v)   the modal properties of *minds* cannot be mind-dependent; thus, (vi) minds cannot be material objects. Therefore,

(vi) naturalists should embrace some form of substance dualism.

The bulk of *World Without Design* is devoted to defending the premises of this argument; and since the defense is both lengthy and complicated, I won't attempt to summarize it here. Rather, I will simply take the conclusion for granted, and focus my attention on the following two questions that were not taken up in the book: First, why does the argument spell trouble for a naturalist? Second, why—as the opening sentence of this paragraph suggests—does the trouble come from the fact that naturalism takes its stand with the metaphysicians rather than the empiricists?

In regards to the first question, the argument spells trouble for a naturalist because, if it is sound, naturalists are committed to believing in things—souls—that cannot be investigated by the methods of science. But, as I indicated earlier, naturalists are united in part by the view that the

correct ontology includes *nothing* that cannot be investigated by the methods of science. Thus we have a point of dissonance. Moreover, the argument depends importantly on the role played by explanatory appeals. The claim that naturalists ought to embrace constructivism is explicitly grounded in the demand for an explanation of our modal knowledge. The claim that constructivism leads to dualism is implicitly grounded in the idea that mind-body dualism provides the best explanation (given constructivism and other constraints imposed by naturalism) for mental phenomena. A van Fraassen-style empiricist might simply beg off of these demands for explanation, but to the extent that naturalists take their stand with the metaphysicians (as characterized above), naturalists cannot dodge the demands. Thus we have our answer to the second question: It is because they take their stand with the metaphysicians that naturalists are forced to accept the untoward ontological consequences that arise out of taking certain demands for explanation seriously.

In sum, then, if the argument just summarized is sound, naturalism falls into dissonance for the following reason: By virtue of its tie to scientific realism, naturalism is committed to taking demands for explanation and inferences to the best explanation with ontological seriousness. But in doing this, it is forced into an ontology that includes things that cannot be investigated by science—an ontology that is different from the sort of ontology to which they take themselves to be committed. Note too that the latter commitment will not be an easy one to give up. It is not as if naturalists *thought* that they were committed to an ontology of atoms but learned *from science* that they were committed, say, to an ontology of fields. Rather, the situation is that, whereas they thought they were committed to a purely scientific ontology, in fact they are forced to postulate entities beyond the reach of science to help explain certain phenomena in the world. Of course, one might respond here by saying that the very fact that souls help to explain phenomena in the world shows that they are not beyond the ken of science after all. But in the mouth of a naturalist this reply can only seem fulsome. For, of course, dualists have always taken souls to be explanatory postulates, and naturalists have tended to insist that souls are inadequate explanatory postulates because, among other things, they are beyond the ken of science.

4. CONCLUSION

My argument in this paper has taken the form of a dilemma: Either naturalism is a thesis, or it is not. If it is a thesis, then it falls into dissonance because dogmatic adherence to a thesis is inconsistent with the naturalistic commitment to follow science where it leads. If it is not a thesis, it is still dissonant, but now for another reason. And the other reason is just this: Naturalism is committed to scientific realism, and also to an ontology that includes only things that can be investigated by science. But the commitment to realism forces naturalists to accept arguments that proceed by way of inference to the best explanation; and one such argument shows that naturalists are committed to substance dualism, a thesis that populates our ontology with entities that cannot be investigated by science. Dissonance then, if the demand for explanation is rejected, and dissonance if it is accepted. Thus, to the extent that a theory is successful only if it avoids falling into dissonance, naturalism is a failure.

5. REFERENCES

Armstrong, D. M. 1980: "Naturalism, Materialism, and First Philosophy". Reprinted in Moser, P. & Trout, J. D. (eds.) *Contemporary Materialism*. New York: Routledge, 35-46.

Forrest, P. 1996: *God Without the Supernatural*: *A Defense of Scientific Theism.* Ithaca: Cornell University Press.

Quine, W. V. O. 1981: *Theories and Things*. Cambridge, MA: Harvard University Press.

Rea, M. 2002: *World Without Design*: *The Ontological Consequences of Naturalism.* Oxford: Clarendon Press.

Schmitt, F. 1995: "Naturalism." In: Kim, J. and Sosa, E. (eds) *Companion to Metaphysics*. Oxford: Blackwell, 343-345.

Sellars, R. W. 1922: *Evolutionary Naturalism.* Chicago: Open Court Publishing Company.

Sellars, W. 1963: "Empiricism and the Philosophy of Mind." In: Sellars, W.: *Science, Perception, and Reality*. London: Routledge and Kegan Paul, 127-196.

Van Fraassen, B. 2002: *The Empirical Stance*. New Haven: Yale University Press.

Naturalized Philosophy of Science: A Cognitive Approach

Ulrich Frey
Technical University of Braunschweig

Current debates in philosophy show an important distinction between naturalistic and non-naturalistic approaches. The explanatory power and shortcomings of both perspectives are discussed, and some consequences of a *naturalistic* perspective in the Philosophy of Science are explored. It is argued that every investigation into the scientific enterprise needs to consider the cognitive abilities of human beings, including scientists. Next to historical and sociological factors cognitive abilities are at the center of problem-solving or decision-making of individual scientists and therefore of paramount importance. Such an empirically based cognitive Philosophy of Science is demonstrated. This has often been called for, but has almost never been implemented. There are hardly any attempts to link empirical results from Cognitive Science to historical scientific case studies. This article argues for such an approach and puts it into practice with three short case studies. The main tenets of a naturalized Philosophy of Science are then derived from these considerations.

1. A NATURALISTIC VIEW OF CURRENT PHILOSOPHY OF SCIENCE

1.1. Empirical Philosophy of Science

This article tries to argue that only a naturalistic approach in the Philosophy of Science is feasible. Three case studies in the history of science are analyzed to strengthen this claim. We will call attention to how often it is necessary to revert to naturalistic means. Any approach to the Philosophy of Science *without* using methodological naturalism, that is results from natural sciences, is unable to explain as much as naturalistic approaches.

This article is not about arguments for or against naturalism. Instead the analysis itself is one long demonstration to show that non-naturalistic approaches have to overcome too many problems, making naturalistic accounts the better analytic tool.

Naturalized Philosophy of Science has its focus on the *cognitive* abilities of scientists. That is to say research has to take into account how people—scientists—think. More precisely: How do they solve problems and which cognitive mechanisms are used for solving the many day-to-day problems? Without considering evolutionary processes this cannot be answered: Like any other sophisticated trait of our body, the brain is an adaptation. Its structure as well as its inherent behavior has been shaped by our evolutionary history. For millions of years our brains had to solve problems of small hunter-and-gatherer groups. It seems evident that this left marks on the way humans think, on what we are able to do and what not. Cognitive abilities are important, not only the often cited historical or social background.

Details about strengths and weaknesses of the human mind can be uncovered by *empirical* investigations, most prominently Cognitive Psychology and Evolutionary Psychology. These disciplines give answers to the *How*? and *Why*? of human cognitive abilities. A complete analysis in Philosophy of Science may not do without integrating empirical results from Cognitive Sciences and Evolutionary Biology. This applies even for *sociological* studies interested for example in group interactions: There are some sociobiological fundamentals (here: about group behavior) which cannot be neglected when analyzing scientists at work. The following cognitive approach takes this into account claiming that there is an important connection between empirical results from evolutionary and cognitive psychology and scientific performance.

To prove that claim we will look more closely at a special case of cognitive abilities: Are there any systematic *scientific errors* directly connected to cognitive errors? If we recognize that the problems of our ancestors *do not match* with the problems of today, *especially* not with scientific problems, then it would appear reasonable to expect that many cognitive errors should arise from that discrepancy. These errors can in turn be explained by Cognitive Science and evolutionary accounts.

1.2. Neglect of cognitive aspects in the Philosophy of Science

Philosophy of Science has been less concerned about cognitive factors so far, rather focusing on historical (Kuhn 1962/1976, Feyerabend 1976) and social aspects (Knorr Cetina 2002) or on rationality and progress (Popper 1934/1994, 1972/1998; Laudan 1977; Lakatos 1978/1982; Kitcher 1993; Callebaut 1993; Vollmer 1993). Exceptions are Tweney et al. (1981), Giere (1988, 1992) and Dunbar (2002), who analyze scientific discoveries from a cognitive-psychological perspective.

In contrast to that noticeable lack of publications in the "Cognitive Philosophy of Science", there are quite a few announcements (see Callebaut 1993; Ruse 1995; Kitcher 1993) demanding to take Cognitive Sciences seriously and incorporate them in case studies under the label *Cognitive Philosophy of Science*—closely linked to Psychology of Science. Unfortunately—and this is an important objection—almost no historical scientific case studies based on empirical results from Cognitive Science are available. There are hardly any attempts to link empirical results from say Biology or Psychology to historical examples.

However, it will not be sufficient to do Cognitive Philosophy of Science. Empirical research from Cognitive Psychology often enough only *describes* human abilities, but does not *explain* them. Possible explanations can be derived from Evolutionary Psychology (Cosmides & Tooby 1992), Evolutionary Epistemology (Vollmer 1975/2002) and Evolutionary Biology. These disciplines rest on a solid empirical basis, which is exactly where philosophers of science should look for support. This article tries to show how such an approach could look like.

In contrast to cognitive aspects, sociological factors influencing scientific work have been studied extensively. In my opinion such kinds of explanation are, however, seriously flawed. I will confine my criticism to three points.

(1) Criticism: "Despite decades of research on this issue, cognitive sociologists have yet to produce a single general law which they are willing to evoke to explain the cognitive fortunes of any scientific theory, from any past period." (Laudan 1977, 217/218, emphasis in the original)

(2) Criticism: If there are indeed negotiations in the laboratory—
and not even the most ardent naturalist would deny that—
what are the negotiations about? Prior in time and logic is
the interaction with entities of the real world (and even
social-constructivists are ontological realists, see Knorr
Cetina in Callebaut 1993).

(3) Criticism: Social-constructivists should be able to show that
identical social circumstances produce in fact the same kind
of science (Kitcher 1993). This has not been shown yet.

2. METHODOLOGICAL NATURALISM AND ITS CRITICS

Before proceeding to the proposed naturalistic (evolutionary-cognitive)
approach it is necessary to define naturalism more precisely.

2.1. A definition of naturalism

The term naturalism is rather vague and often used in many different
notions by proponents and opponents alike. In its most trivial sense it
means taking natural sciences seriously—a statement everybody would
agree with. A more ambitious and precise definition qualifies naturalism in
ontological respects as materialistic, in *methodological* respects as making
the heaviest possible use of natural sciences and in *epistemological*
respects as proposing a hypothetical realism (see Sukopp 2006).

This article will mostly be concerned with methodological aspects, that
is how Philosophy of Science relates to natural sciences. It depends on
natural sciences in at least three ways. Firstly, and rather trivially so,
science and natural sciences in particular are its *subject*. Secondly,
philosophers have to use data from natural sciences to support their
hypotheses and claims—there is no such thing as *a priori* philosophy.
Some might raise the objection that this is a circular argument as the
studied subject itself is used to support theories about it; but this is
common for many sciences: language is used to do linguistic studies,
neurophysiologists use their brains to research them, theories about groups
are discussed in groups, and so on. These examples are *virtuous* circles
(see Vollmer 1975, 113f.). Thirdly, the hypotheses about science

themselves have to be grounded in empirical research. If we suppose that a theory about the influence of scientific group structure on its research output is proposed, this claim would certainly have to recognize some very basic facts about groups known from sociobiology. If the claim contradicts well-established sociobiological facts, then in most cases the hypothesis should deservedly be rejected.

Furthermore, it cannot be denied that modern science contributes to genuine philosophical questions. Take Philosophy of Mind. Spectacular details from PET-scans and fMRI (functional Magnetic Resonance Imaging) allow new insights and are about to explain ancient riddles of Philosophy like freedom of will, body-mind problem etc. This, however, is not undisputed; the common reaction from *philosophers* is to cede to the empirical results while declaring that the problem itself is not touched by it. This I would call a forced withdrawal.

Some areas of research within the natural sciences, however, *are essential* for Philosophy of Science. Among them is Evolutionary Biology answering questions about our origin and our biological history as well as Cognitive Psychology answering questions about our mental abilities. Other sub-disciplines like Evolutionary Psychology or Sociobiology could be included making the relevant subjects by no means complete. The next section is dedicated to defending these essential disciplines against criticism.

2.2. Rejection of criticism on Evolutionary Epistemology and Evolutionary Psychology

The most common (but not the most valid) criticism against evolutionary accounts, such as Evolutionary Epistemology or Evolutionary Psychology, include the criticism of *nativism*, *panadaptionism* and the claim that evolutionary theories are *not falsifiable*. Neither of the first two are in fact claimed by any proponent (see for example Cosmides and Tooby 1994), while the third point cannot really be doubted: There are numerous empirical studies with over 10.000 humans which *do* make testable and falsifiable predictions (see Buss 1989).

Some other criticism—like the weak empirical basis of Evolutionary Psychology—looses rapidly justification: Recent neurophysiology results

and empirical psychology support its predictions very well. The most fundamental critique claims that Evolutionary Epistemology is not a "complete" epistemology; real philosophical problems are not within its explanatory framework. However, the first is true for all epistemological approaches (*tu quoque*-argument). Moreover, philosophers are well advised to *confine* their theories, as too encompassing approaches have regularly turned out to be too ambitious. In fact, approaches including Evolutionary Epistemology which cover "only" part of the whole usually turn out to be both more useful and more correct than "big picture" theories.

   Summing it up: It is safe to say that disciplines based on evolutionary theory both stood the test of time and defended themselves well against criticism—they are testable, empirically precise, and conceptually valid. It is on the contrary approaches which *do not* take evolutionary aspects into account that have to justify their applicability and claims.

2.3. Rejection of criticism on Cognitive Philosophy of Science

As this article voices claims about cognitive factors in science these hypotheses shall be defended against criticism, too.

   (1) *Criticism*: Cognitive Philosophy of Science has no *philosophical relevance*, it does not solve any philosophical problem.
        Reply: It is true that this approach is an example of a *descriptive*, *concrete* and *empirical oriented* Philosophy of Science. It is the opposite of the typical, often found very abstract, broad and general *top-down*-approach. Especially *because* such speculative approaches failed so often in the past when applied to specific historical examples, this provides an argument of how *not* to do it.
   (2) *Criticism*: In order to investigate science (being a cultural activity) evolutionary shaped cognitive thinking patterns are not relevant at all. The connection is too far-fetched or even non-existent.

Reply: That is the very connection this article tries to show. The "gap" can in fact be closed quite easily (see section 3). Cognitive peculiarities of humans—in this case, cognitive errors—occur both *systematically* and *repeatedly* in science. These errors are structurally *identical* with cognitive errors produced in controlled experiments in the laboratory. The easiest and most obvious explanation for these errors is that they originate in the *same source*, for example a certain error-prone thinking pattern. It will be shown that this *relates* the two levels seemingly so far apart.

(3) *Criticism*: There are other, better explanations. Theories and their errors are dependent on the *background theories* of each era and are adapted accordingly to the changing background.

Reply: Historical examples show that this hypothesis is not entirely tenable. It cannot explain why the same errors occur in medicine in the 16th, in physics in the 18th as well as in psychology in the 20th century, which can be shown. Practically every part of the so called "background" has been changed or has never been the same in the first place.

3. A NATURALISTIC ACCOUNT OF SCIENTIFIC ERRORS

So far the article has argued against possible criticism. Now, positive evidence for a Cognitive Philosophy of Science shall be presented. The focus is on cognitive errors of scientists, a research topic very suitable to show that Philosophy of Science has to rely on natural sciences when trying to explain phenomena from the history of science. The first example is the framing effect in medicine.

3.1. The framing effect in medicine

Physicians sometimes tend to have "peculiar" habits when deciding about therapies. If they heard or read about a therapy in a *negative* formulation, like "7 out of 100 patients die from this treatment" they are more likely to abandon this kind of therapy than when they encounter the same therapy in

a *positive* formulation, like "93 out of 100 patients survive this treatment". This is called the *framing effect* (for an overview see McGettigan et al. 1999).

This decision making procedure is remarkable, as the *same* number of patients will die or survive in both formulations. The only difference is the *frame*—negative vs. positive. But surely this simple "trick" cannot influence trained experts, physicians with a doctoral degree and years of practical experience? Surprisingly, however, it can, and the effects are quite large: 25 to 42 per cent are influenced by the framing effect (McNeil et al. 1982). How are we to explain this fact? Why do scientists (here: physicians) behave so irrationally? While traditional philosophers of science are at a loss to explain, this is nevertheless a valid question. In contrast, for naturalized philosophers of science the answer is easy, as the framing effect is a rather well-known fact in Cognitive Psychology. Studies concerning the framing effect comprise well over 30.000 people of all ages, professions and education levels (see Kühberger 1998). It is simply one of the many cognitive errors human beings commit. It is trivial, yet has dramatic implications for the patients. Additionally, it is systematic. Detailed empirical studies show how and when this phenomenon occurs and supply the data including an explanation.

This kind of naturalistic analysis—coupling cognitive errors known from the natural sciences with (historical) scientific case studies—is obviously a powerful method. It is surprising how often errors in science have *already* been described in natural sciences, such as Cognitive Psychology and this can be backed up by further examples coming from widely different fields.

3.2. Common errors with complex systems

Many experiments have shown that the majority of subjects (laymen) have great difficulties with complex systems. However, almost every scientific problem is in some form complex, which means that errors should be common in problem-solving strategies of scientists, always assuming that scientists are not totally different from laymen at least in cognitive respects. A closer look reveals in fact that some errors occur again and

again when subjects handle complex systems (after Dörner 1989). The most common are:

(1) Oversimplification (complex systems are reduced to simple, linear systems with *one* dominant cause. Side effects, long range effects, feedback-loops and other decisive characteristics of the system are simply ignored.).

(2) Overgeneralization (a general rule is deduced from one or few cases).

(3) Underestimation (a general tendency to underestimate the overall complexity).

(4) Isolated treatment (networks are treated as isolated systems).

(5) Underestimation of non-linear processes.

(6) Fixation on the situation (the momentary situation is regulated instead of the dynamic process).

(7) Too little control (no checks, too little gathering of information, no analysis of own measures).

(8) Blindness against own errors.

(9) Not learning from own errors.

(10) Wrong problem-solving priorities (problems are solved by conspicuousness instead of importance).

(11) Misplaced sphere of competence (methods of one's own competence are transferred to areas unsuitable for them, thus producing errors).

These are just some of the most common reactions and strategies of most humans when they have to deal with complex systems. We can classify these errors in three categories: Firstly, *oversimplification* on all levels (1-4). Secondly, significant problems with *time* (5-6). Thirdly, errors in *strategy-choice* (6-11).

The most severe error is oversimplification, that is reduction of complex systems (1-4). In most cases a *linear* problem-solving strategy is used: Problem A has to be solved, therefore solution A'. This, however, creates problem B; this is solved, but creates problem C and so on. This cascade of problems could have been avoided in the first place, if B and C had been anticipated while solving A; However, this is practically never the case.

This short and necessarily incomplete account of problems with complex systems should make us wary whenever we watch scientists work with complex systems. It is not unlikely that they will commit the same errors as laymen. Three case studies (section 4) will demonstrate that this is indeed the case. But still an *explanation* is needed why human beings have so much trouble handling complex systems; so far, cognitive errors have only been *described*.

3.3. Evolutionary explanation

The way in which human beings handle complex situations is a reaction to a certain environment. It is a *complex* world we live in. It must be *processed* so that we can handle it. In a lot of situations it is imperative to act *fast* while at the same time we must not be impaired by *incomplete* or *wrong information* in our decisions.

All the said requirements are fulfilled by the above mentioned "errors". In a more positive formulation these *mechanisms* enable us to act at all: Complicated estimates are reduced to simple and linear extrapolations, networks of multiple causes and effects are simplified to one cause, alternatives are cut off, measures taken are not controlled and just a few examples are sufficient to deduce general laws.

From an evolutionary perspective these "errors" are very appropriate measures to come to terms with complex situations under time pressure, incomplete information and capacity limitations (for example memory). To achieve this, humans use heuristics that succeed in trading off minimal resources with sufficient precision. These adaptations aim at *swiftness* and *simplicity*. This is predominantly achieved by reduction as is evident from the mechanisms. A short look at our environment is enough to show why this *has to be* like that: It is impossible to process all the information. Only 1 % of the *already strongly filtered* sensory input is computed, much less is actually used or remembered (von Ditfurth 1976). Furthermore there are nearly endless combinatorial possibilities of the existing data. This requires radical reduction and radical simplification on all levels. This has been spelled out in detail and tested empirically (see Dörner 1989; Gigerenzer & Todd 1999).

4. HISTORICAL SCIENTIFIC CASE STUDIES

It should have become clear by now that natural sciences like Cognitive Psychology and Evolutionary Biology can offer good *descriptions* and *explanations* of phenomena that are of interest in the Philosophy of Science. The fact that cognitive explanations are indeed *at the very center* of research is demonstrated in the following three historical scientific case studies where exactly the same cognitive errors are found as known from experiments in the laboratory.

4.1. Case study 1: Introduction of new species—snails

Ecosystems are complex systems. Humans have tampered with ecosystems as long as human history, due to which many species have become extinct (for example the Dodo or the Cassowary). However, *introductions* of species in existing ecosystems are every bit as disastrous. The following examples deal only with *deliberate* introductions by scientists or by politicians guided by scientific advice, to make sure that only scientific errors and not political considerations are investigated.

A very prominent example of *linear problem-solving* is the attempt to repair former mistakes: If an introduced organism does not fulfill its assigned role, the typical reaction is to introduce *another* organism. This second organism, so the argument, shall fight and destroy the first one.

One first example is the giant snail *Achatina fulica*, which was introduced in Asia as food source for the local population around 1910 to 1940, in Hawaii in 1955 (Cook 1989) and in Tahiti in 1967 (Murray et al. 1988). Growing exponentially, *Achatina* rapidly became the worst snail pest in the tropics (Mead 1961; Cowie 2001). For that reason two predatory snail species, *Euglandina rosea* and *Gonaxis kibweziensis* were introduced in Hawaii. The first did not only fail to prey on *Achatina* as expected (Griffiths et al. 1993, 79), but exterminated nearly every endemic snail species (Cowie 2000, 145)! Only 10 to 35 % of the former 750 species are alive today (Cowie 2001). The same error was made when many *Partula*-snails in Tahiti went extinct by *Euglandina* (Cook 1989). In the mentioned cases the decision to introduce the predatory snails was reached by committees of scientific experts (see Mead 1961, 128ff). And

even after *Euglandina* another twelve snails have been introduced with the same intent but without success (Cowie 2001).

Many more errors are apparent: The ecosystem as a *network* is not considered at all (instead: two snails, nothing else), feedback-loops and interdependencies are neglected (the fact that *Euglandina* feeds on other snails). Instead we notice a radical *simplification* towards *one* goal (extinction of *Achatina*), which shall be accomplished by *one* measure alone (*Euglandina*). If the first "solution" does not work, then the same strategy is *repeated* with even worse consequences (*Euglandina* and then twelve other snails).

Moreover, in many countries the introductions were never tested before their implementation. There was neither a *control* of the measures nor an *analysis of errors* (a typical error, see section 3.2). The one study in Hawaii that was done before *Euglandina* was introduced consisted of determining whether *Euglandina* preyed upon *Achatina* at all. It did, although once in the wild it unfortunately preferred nearly entirely smaller snails (Cook 1989). And worse: as soon as *Euglandina* was introduced into the wild, nobody cared about what happened at all.

> However, the lack of follow-up, both in terms of determining its impact on crop pests and in ensuring that it does not cause damage to native species, is striking. (Griffiths et al. 1993, 80)

There was neither a detailed search for information nor reliable empirical data about how *Euglandina* and *Achatina* populations would interact when in the same habitat (Christensen 1984; Murray et al. 1988). On the contrary, there was convincing evidence that they would *not* interact (see Cowie 2001). Even the commissioned scientific expert for that purpose, Kondo, and the expert on *Achatina*, Mead, had to confess their ignorance—but recommended the introduction nevertheless, although Mead *himself* cites half a page of literature indicating the dangers of new introductions (Mead 1961). But as soon as 1957 it was known by experiments that *Euglandina* preferred smaller snails to *Achatina*—even if experts still predicted the opposite (Mead 1961, 132).

These examples show the transferability of many cognitive errors from experimental test situations in the laboratory to actual decisions by

scientists—they are in fact committed by scientists, experts, and laymen *alike*.

4.2. Case study 2: More harmful introductions of species

Psychological research (Dörner 1989, see section 3.2) shows that it is very hard to learn from own mistakes and this is exactly what can be seen in many historical examples encompassing disciplines, countries, and epochs:

> A major theme of this book [about harmful introductions] is that we do not learn from history, that we continue to emulate the mistakes of the past. (Low 1999, XXVI)

Interestingly enough, the success quota of biological introductions to fight other introduced species is well-known: Only 6 % were completely successful, 18 % were somewhat positive, but an overwhelming 76 % were totally unsuccessful (Low 1999; see also Cowie 2001 for similar figures for snail introductions).

In spite of that there are plenty of current examples: As late as 1988, Australia introduced a grass (*Hymenachne amplexicaulis*) despite being very aware of negative effects of biological introductions. Only eleven years later this grass is counted as one of the top twenty worst weeds, because it displaces other endemic species. Two years later, in 1990, Kochia (*Bassia scoparia*) was introduced for land reclamation; only five years later half a million Dollar had to be invested to get the problem under control. The Australian government based both decisions on scientific expertise and was—due to former unsuccessful introductions—clearly aware of harmful effects accompanying introductions of new species; thus, these considerations played a central part in the decision-making process. Nevertheless the errors were committed again. For an overview of harmful introductions, see Kowarik (2003) and the *Invasive Species Specialist Group* (Invasive Species Specialist Group (ISSG) 2005).

Although most of the introductions of the past were quite problematic or even bordering on catastrophes, the future will be even worse: Genetically altered organisms have a still higher destructive potential. The consequences of such introductions are not in the least foreseeable, much less controllable. So even right now we do not learn from well-known

mistakes in the past, but instead increase the potential for disaster at a scale not yet to be overseen: The genetically altered seaweed *Caulerpa taxifolia* is seven times bigger as its unaltered counterpart and has escaped to the Mediterranean Sea beginning to cover increasingly larger parts of the sea floor (Low 1999).

At the same time this last case is an example for another error—to have too little information to act in an appropriate manner but act nevertheless. The same goes for the introduction of the Aga-toad (*Bufo marinus*) in Australia to fight a pest feeding on sugar cane—the only information about the suitability of the toad to fight the pest was the correlation of the population sizes of these two species in *one* country (Puerto Rico) in *one* year (1931). There was indeed a decline of the pest in this year—however, this was due to high rainfalls, not to the toad. Today the toad is one of the most devastating pests in Australia (Low 1999).

Adding to this underestimation of complexity and the resulting insufficient gathering of information are three more errors: Future trends are mostly extrapolated in a *linear* way, long-range effects are *neglected* and cause and effects with large time-lags between the two are extremely difficult to master:

> Another difficulty is that cause and effect, in invasion biology, are often far removed from each other, separated by time, and often by space. (Low 1999, 293)

Many more cognitive errors of scientists could be demonstrated (see Frey 2007), but the similarity of errors known from the laboratory and historical case studies should be apparent by now.

4.3. Case study 3: Long-term management of ecosystems

The best examples to illustrate the general inability of humans to manage complex systems are those where *the opposite* of the desired and planned effects occurred. Take the management of the Blue Mountain forests in Oregon, USA. Over a 100-year period (1900-2006) we see case after case of mismanagement:

> In the process of trying to manage extremely complex landscapes, foresters set into motion a chain of events that increasingly swung out of control. (Langston 1995, 8)

The most important goal of the National Forest Service—since 1900—has been to regulate the forest in order to create sustainable, "efficient", and productive forests consisting mostly of valuable pines. A look at the forest today shows us that *not a single one* of these most important goals has been achieved. Instead, the opposite developments emerged.

The first example are firs. After 100 years of efforts to increase the pine population *firs* have become the most prevalent tree in the forest and have replaced the desired pines for the most part. This in turn has had unwanted and unforeseen *side-effects*. One of them is the massive loss through pests against which firs are less resistant than pines. Although this fact has been known since 1913 (see Bright 1913) it came as a surprise to foresters.

Why did firs replace pines? It was known that pines needed more light than firs. Therefore, cutting more clearings was thought to be a sufficient measure. This *oversimplification* on one aspect of the problem is typical for linear problem-solving. Foresters missed another crucial factor—firs grow faster than pines, thus the clearings were colonized by firs (Langston 1995; the preferred cutting of pines is another contributing reason). As the pines did not grow as anticipated by this measure this approach was abandoned altogether—there was no more sowing of pine seeds *at all*. Instead a new *single* solution was proposed: proper fire fighting. Unfortunately this, too, did not work out as planned.

A second example is the fire fighting itself. To foresters "efficient" management meant to fight small and middle-sized fires. This, however, lead to high losses through uncontrollable major fires. These major fires occurred only because fuels could accumulate in great amounts *because* small fires were extinguished:

> The very effort to avoid forest fires has helped, in a later generation, to create them. (Little 1995, 84)

Moreover, these misguided attempts of fire fighting made the pest problem much worse:

> […] a kind of worst-case scenario come true. (Little 1995, 84)

Nevertheless the National Forest Service does not appear to have learnt from its own mistakes: The new fire fighting program costs are two billion US-Dollar just for the year 2000 (USDA Forest Service 2005a), although it is known from the management of the Yellowstone National Park that the most efficient and least expensive fire fighting is: doing nothing at all. Almost all fires burn out relatively quickly by themselves (Clark & Minta 1994).

We notice at least three unwanted results of the forest regulation: firs instead of pines, devastating fires and pest problems. Let's take a closer look at the measures that produced the pest problems:

The massive pest problem began around 1969. Two main causes can be identified: Firs were much more susceptible to pests than pines and dead wood was lacking, because it had been cleared away to increase productivity. Dead wood, however, is an integral part of forests, providing nutrients, shadow and humidity for young trees. Furthermore it is the preferred habitat of predatory insects and insect-eating ants (*Campotonus*). These ants in turn are the main food source (98 %) of woodpeckers. Woodpeckers, however, were neglected in considerations, because they do not prey on pests. *But* they create nesting opportunities for other birds preying on insects. As neither ants nor birds controlled them, the number of insects exploded. Very fast this turned into a catastrophe: In 1990 50 % of the Douglas-firs had been destroyed by insects and 63 % of all trees had been damaged (Langston 1995; Little 1995).

If human beings were good at learning from their mistakes, this could be corrected in the future. Unfortunately, this is not the case, as can be seen in the current forest plan (valid until 2006): Again the National Forest Service expects maximum growth, coming about through a disappearance of all limiting factors (for example no insects) coupled with an increase of all positive factors. The enormous conflicts between grazing, protection of biodiversity, forest growth and the recreational value of the forest are mentioned, but "solved" by referring to scientific progress. This progress will make it possible to attain two goals at the same time, even if they are mutually exclusive (USDA Forest Service 1990, 3f. to 4-17). These errors are by now familiar to us, as they occur again and again and are well-known from controlled laboratory experiments.

These three examples (5.1, 5.2, 5.3) have shown how a cognitive analysis of historical case studies could look like. The next two chapters convert this analysis into arguments for naturalistic approaches in general.

5. ARGUMENTS FOR AN EVOLUTIONARY-COGNITIVE PHILOSOPHY OF SCIENCE

The probably most convincing arguments for an evolutionary-cognitive approach, that is the systematic integration of experimental evidence from natural sciences into Philosophy of Science are as follows:

(1) *Argument of cognition*: It can't be doubted that science is primarily a cognitive activity. The basic level of science is about daily problem-solving, mental model building, finding new analogies, and so on. Generally speaking, how scientists *think* should be of high importance to philosophers of science (and, of course, to scientists themselves).

(2) *Argument of relevance*: The relevance of psychological experiments concerning thinking is high and hardly disputable. On the one hand there are studies in the Philosophy of Science connecting these two convincingly (for example Wimsatt 1980). On the other hand independent historical studies which have nothing to do with Psychology describe strategies and errors of scientists as if they were describing the cognitive errors known from experimental psychology—*without knowing that the latter exist*. Complementary, there are psychological studies which establish a direct link from psychological errors to errors in science (see section 3.1; McNeil et al. 1982; Wimsatt 1980).

(3) *Argument of Similarity*: The similarity of experimentally proven errors with existing errors in the history of science is remarkable. It is often possible to see them as identical. There is no reason to pass over the most simple and most evident explanation: These errors are based on the same faulty cognitive mechanisms.

(4) *Argument of invariance*: Many historical examples show clearly the *invariance* of errors. This is important, because it excludes rival explanations: A look at the "cold fusion" episode in 1989 shows that many errors—excluding fraud issues—occurred. But *paradigmatic* interpretations have a hard time explaining them, as there was no paradigm of "cold fusion". A similar problem arises for *historical* interpretations for very similar, if not identical phenomena occurring in very different epochs. One salient example are the attempts to measure skulls and intelligence differences between races (Gould 1988). It is easy to find very similar errors from 1840 to 2006, while the historical preconditions and background theories vary considerably during that time.

(5) *Argument of being the only explanation*: For the case study of handling complex systems (here: the management of ecosystems) the evolutionary-cognitive explanation is the *only* valid one. For how else can one hope to understand the apparent and systematic inability of humans to handle complex systems? It seems evident that humans are simply not made for this and therefore reach their cognitive limits.

(6) *Argument of universality*: The best argument for evolutionary-cognitive explanations is evidence from different cultures and (if possible) from children. This evidence can in fact be provided (see Samuels & McDonald 2002; Cosmides & Tooby 1997). Furthermore, many historical examples span centuries as well as disciplines as different as anthropology, physics, medicine or psychology (see Frey 2007). Further evidence comes from Gigerenzer and Hoffrage (1995): The more ecologically valid (that is similar to natural environments for humans) tasks are formulated, the better the subjects perform.

These arguments are at the heart of a Cognitive Philosophy of Science. But cognitive factors are certainly not the only influencing element. There are many interactions between cognitive, social and historical influences which weaken and amplify each other.

After these six arguments in favor of my hypothesis, two possible falsifications shall be mentioned: My hypothesis would be falsified, if an independent historical examination of the used case studies cannot find the described errors or if many new case studies were completely without evolutionary-cognitive aspects.

6. ARGUMENTS FOR A NATURALIZED PHILOSOPHY OF SCIENCE

The data mentioned above shall now be integrated into a naturalized Philosophy of Science. Often enough, proposals on how to do Philosophy of Science have a weakness: They do not describe the real history of science accurately enough and too often have not even tried to. If—and this is true even for the most famous philosophers such as Kuhn or Lakatos—their theoretical postulates are *tested empirically* with diverse historical case studies, most postulates turn out to be outright false (see Donovan, Laudan & Laudan 1988/1992).

One of the most promising attempts to solve these intricate problems is that of a naturalized Philosophy of Science, as advocated by Giere (1988; 1992). This approach in its minimal form has two demands: The first methodological rule is that empirical evidence is of foremost importance. Each and every theory has to be verified by case studies. Secondly, results from natural sciences have to be at the core of Philosophy of Science: This applies to generating hypotheses as well as to testing them. Both demands have not been met in Philosophy of Science, although this has been slowly changing in the last years.

In this article I have tried to show the importance of these two points. As can be seen by the three short case studies it is indispensable to rely on results from natural sciences, and the two disciplines with probably the most prominent influence on scientific activities are Cognitive Psychology and Evolutionary Biology.

Take one example from the very heart of Philosophy of Science: the principle of falsification. This is, since Popper, one of the most important tools for evaluating theories. But *are* human beings able to falsify correctly and do they in practice try to falsify their own theories? The sobering answer is: no. Wason (1968) and many follow-up studies (for example Griggs & Cox 1982) show that only about 2-5 % of the tested persons

solve an easy task by using falsification. So, although the theoretical necessity to falsify remains unchallenged, falsification in practice is far from being used consistently.

The necessity of an integrated, naturalized analysis is also evident from the historical development of Philosophy of Science in the 20th century. All modern conceptions differentiated science more and more, painting a picture where scientists are under the spell of *many* influences, determining their work. All these descriptions, however, are too narrow: they confine themselves to social and historical influences, while other influences have almost completely been forgotten: Human beings are not only socio-cultural beings, but also biological organisms. This is the point where a naturalized Philosophy of Science has its focus and complements traditional approaches. Psychological and biological aspects of our mental structure have to be considered.

Philosophy of Science (and Epistemology) is researching an empirical phenomenon, that is the cognitive products and methods of biological human subjects with certain abilities as well as limits. This puts cognitive products at the center of attention. If one sees science primarily as a problem-solving activity—and many philosophers of science do that (for example Popper 1972/1998; Lakatos 1978/1982; Laudan 1977)—then the focus of research has to be the process of thinking, more precisely the process of problem-solving. Thus, the level of our inquiry refers to the *daily* scientific work which consists of a huge number of small decisions, evaluations, solutions to practical problems and so on. But many philosophers of science choose a *coarser* unit of inquiry, for example *research programs* or *research traditions*. Many phenomena are missed that way, because strengths and weaknesses of our thinking processes have to be considered for answering them.

It is difficult to say why this important approach has been overlooked for so long. One possible reason is that it seems impossible to generalize individual cognitive processes. However, naturalists are only interested in the cognitive processes that are *common to all individuals*. There are loads of data concerning human problem-solving, decision-theory, confirmation of hypotheses which are all relevant to discoveries (for example Tooby & Cosmides 1992; Gigerenzer 1999).

7. CONCLUDING REMARKS

I have tried to emphasize how essential it is for Philosophy of Science to rely on natural sciences as methodological means. However, there are points where naturalists can agree with non-naturalists. I agree with Kuhn, Feyerabend, Latour and Knorr Cetina in their criticism of the "received view" of science as being too simplistic. Social and historical factors *are* important. It is, however, indispensable to add another important influencing factor: the cognitive side of science (psychological and biological explanations). An account focusing on cognitive aspects can *complement* these theories. This has been an area of neglect to date, but humans are first and foremost *biological* beings.

To sum up: What can one expect of a naturalistic approach to the Philosophy of Science?

(1) The focus is on the cognitive (psychological and biological) attributes of human beings (scientists).
(2) These attributes are not subject to speculation, as they are supported by many empirical results from Cognitive Science.
(3) These results in turn can be explained by Evolutionary Biology and Evolutionary Psychology.
(4) Historical and social factors are not neglected—they are complemented.
(5) Theories in the Philosophy of Science need always be warranted by historical case studies as another source of empirical support.

With these points in mind modern Philosophy of Science should be able to progress in describing science more and more precisely—a naturalized Philosophy of Science *is* progressive.

8. REFERENCES

Bright, G. A. 1913: *Umatilla National Forest Annual Silvical Report*. The relative merits of western larch and Douglas-fir in the Blue Mountains, Oregon.
http://www.fs.fed.us/r6/uma/history/index.htm (20.02.06).

Buss, D. M. 1989: Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences* 12, 1-49.

Callebaut, W. 1993: "Conversations with William Bechtel, Robert N. Brandon, Richard M. Burian, Donald T. Campbell, (...)." In Callebaut, W. (ed.) *Taking the Naturalistic Turn or How Real Philosophy of Science is Done*. Chicago: University of Chicago Press.

Christensen, C. C. 1984: Are Euglandina and Gonaxis Effective Agents for Biological Control of the Giant African Snail in Hawaii? *American Malacological Bulletin* 2, 98-99.

Clark, T. W. and Minta, S. C. 1994: *Greater Yellowstone's Future: Prospects for Ecosystem Science, Management, and Policy*. Moose, Wyoming: Homestead Publishing.

Cook, A. 1989: "Factors Affecting Prey choice and Feeding Technique in the Carnivorous Snail *EUGLANDINA ROSEA* FERUSSAC." *Journal of Molluscan Studies* 55, 469-477.

Cosmides, L. and Tooby, J. 1994: Beyond intuition and instinct blindness: toward an evolutionarily rigorous cognitive science. *Cognition* 50, 41-77.

Cosmides, L./Tooby, J. 1997: *Evolutionary Psychology: A Primer*.
http://cogweb.ucla.edu/ep/EP-primer.html (14.07.05).

Cowie, R. H. 2000: "Non-indigenous land and freshwater molluscs in the islands of the Pacific: Conservation impacts and threats." In: Sherley, G. (ed.) *Invasive species in the Pacific: A technical review and regional strategy*. Apia, Samoa: South Pacific Regional Environment Program, 143-172.

Cowie, R. H. 2001: "Can snails ever be effective and safe biocontrol agents?" *International Journal of Pest Management* 47, 23-40.

Donovan, A., Laudan, L. and Laudan, R. 1988: "Testing Theories of Scientific Change." In: Donovan, A., Laudan, L. and Laudan, R. (eds.) *Scrutinizing Science: Empirical Studies of Scientific Change*. Baltimore: John Hopkins University Press, 3-44.

Dörner, D. 1989: *Die Logik des Mißlingens: Strategisches Denken in komplexen Situationen*. Reinbek bei Hamburg: Rowohlt.

Dunbar, K. N. 2002: "Understanding the role of cognition in science: the Science as Category framework." In: Carruthers, P., Stich, S. and Siegal, M. (eds.) *The cognitive basis of science*. Cambridge: Cambridge University Press, 154-170.

Federspiel, K. and Herbst, V. 1991/1996: *Stiftung Warentest: Die andere Medizin: Nutzen und Risiken sanfter Heilmethoden*. Braunschweig: Westermann.

Feyerabend, P. K. 1976: *Wider den Methodenzwang: Skizze einer anarchistischen Erkenntnistheorie*. Frankfurt am Main: Suhrkamp.

Frey, U. 2007: *Der blinde Fleck – Kognitive Fehler in der Naturwissenschaft und ihre evolutionsbiologischen Grundlagen*. Frankfurt: Ontos.

Giere, R. N. 1988: *Explaining Science: A Cognitive Approach*. Chicago: University of Chicago Press.

Giere, R. N. 1992: "Introduction: Cognitive Models of Science." In Giere, R. N. and Feigl, H. (eds.) *Cognitive Models of Science*. Minnesota Studies in the Philosophy of Science Volume XV. Minneapolis: University of Minnesota Press, xv-xxviii.

Gigerenzer, G. and Todd, P. M. 1999: Fast and Frugal Heuristics: The Adaptive Toolbox. In Gigerenzer, G., Todd, P. M. and ABC Research Group (eds.) *Simple heuristics that make us smart*. New York: Oxford University Press, 3-34.

Gigerenzer, G. and Hoffrage, U. 1995: "How to Improve Bayesian Reasoning without Instruction: Frequency Formats." *Psychological Review* 102, 684-704.

Gould, S. J. 1988: *Der falsch vermessene Mensch*. Frankfurt am Main: Suhrkamp.

Griffiths, O., Cook, A. and Wells, S. M. 1993: "The diet of the introduced carnivorous snail Euglandina rosea in Mauritius and its implications

for threatened island gastropod faunas." *Journal of the Zoological Society of London* 229, 79-89.

Griggs, R. A. and Cox, J. R. 1982: "The elusive thematic-materials effect in Wason's selection Task." *British Journal of Psychology* 73, 407-420.

Invasive Species Specialist Group (ISSG) (2005): *100 of the World's Worst Invasive Alien Species.*
http://www.issg.org/database/species/search.asp?st=100ss&fr=1&sts=sss (31.01.06).

Kitcher, P. 1993: *The Advancement of Science: Science without Legend, Objectivity without Illusions.* New York: Oxford University Press.

Knorr Cetina, K. 2002: *Wissenskulturen: Ein Vergleich naturwissenschaftlicher Wissensformen.* Frankfurt am Main: Suhrkamp.

Kowarik, I. 2003: *Biologische Invasionen: Neophyten und Neozoen in Mitteleuropa.* Stuttgart: Ulmer.

Kühberger, A. 1998: "The Influence of Framing on Risky Decisions: A Meta-analysis." *Organizational Behaviour and Human Decision Processes* 75, 23-55.

Kuhn, T. S. 1962/1976: *Die Struktur wissenschaftlicher Revolutionen.* Frankfurt am Main: Suhrkamp.

Lakatos, I. 1978/1982: *Die Methodologie der wissenschaftlichen Forschungsprogramme*: *Philosophische Schriften Band 1*. Edited by Worrall, J. and Currie, G. Braunschweig: Vieweg.

Langston, N. 1995: *Forest Dreams, Forest Nightmares: The Paradox of Old Growth in the Inland West.* Seattle: University of Washington Press.

Laudan, L. 1977: *Progress and its Problems.* Berkeley: University of California Press.

Little, C. E. 1995: *The Dying of the Trees: The Pandemic in America's Forests.* New York: Viking Penguin.

Low, T. 1999: *Feral Future: The Untold Story of Australia's Exotic Invaders.* Chicago: University of Chicago Press.

McGettigan, P., Sly, K., O'Connell, D., Hill, S. and Henry, D. 1999: "The Effects of Information Framing on the Practices of Physicians." *Journal of General Internal Medicine* 14, 633-642.

McNeil, B. J., Pauker, S. G., Sox, H. C. and Tversky, A. 1982: "On the elicitation of preferences for alternative therapies." *New England Journal of Medicine* 306, 1259-1262.

Mead, A. R. 1961: *The Giant African Snail: A Problem in Economic Malacology*. Chicago: University of Chicago Press.

Murray, J., Murray, E., Johnson, M. S. and Clarke, B. 1988: "The Extinction of *Partula* on Moorea." *Pacific Science* 42, 150-153.

Popper, K. 1972/1998: *Objektive Erkenntnis: Ein evolutionärer Entwurf.* Hamburg: Hoffmann und Campe.

Popper, K. R. 1934/1994: *Logik der Forschung*. Tübingen: Mohr Siebeck.

Ruse, M. 1995: *Evolutionary Naturalism*: *Selected Essays*. London: Routledge.

Samuels, M. C. and McDonald, J. 2002: "Elementary School-Age Children's Capacity to Choose Positive Diagnostic and Negative Diagnostic Tests." *Child Development* 73, 857-866.

Sukopp, T. 2006: *Naturalismus – Kritik und Verteidigung erkenntnistheoretischer Positionen*. Frankfurt am Main: Ontos.

Tweney, R. D., Doherty, M. E. and Mynatt, C. R. 1981: "General Introduction." In: Tweney, R. D., Doherty, M. E. and Mynatt, C. R.: *On Scientific Thinking*. New York: Columbia University Press.

USDA Forest Service 1990: *Umatilla Forest Land and Resource Management Plan.*
http://www.fs.fed.us/r6/uma/projects/index.shtml (15.02.06).

USDA Forest Service 2005: *Research & Development 2004 Annual Report.*
http://www.fs.fed.us/research/pdf/2004-annual-report.pdf (10.03.06).

Vollmer, G. 1975/2002: *Evolutionäre Erkenntnistheorie: Angeborene Erkenntnisstrukturen im Kontext von Biologie, Psychologie, Linguistik, Philosophie und Wissenschaftstheorie*. Stuttgart: Hirzel.

Vollmer, G. 1985/2003: *Was können wir wissen?* Band 1. Stuttgart: Hirzel.

Vollmer, G. 1986/2003: *Was können wir wissen?* Band 2. Stuttgart: Hirzel.

Vollmer, G. 1993: *Wissenschaftstheorie im Einsatz: Beiträge zu einer selbstkritischen Wissenschaftstheorie*. Stuttgart: Hirzel.

Von Ditfurth, H. 1976: *Der Geist fiel nicht vom Himmel: Die Evolution unseres Bewußtseins*. Hamburg: Hoffmann und Campe.

Wason, P. 1968: "Reasoning about a Rule." *Quarterly Journal of Experimental Psychology* 20, 273-281.

Wimsatt, W. C. 1980: "Reductionist research strategies and their biases in the units of selection controversy." In Nickles, T. (ed.): *Scientific Discovery: Case Studies*. Dordrecht: Reidel, 213-259.

Passing by the Naturalistic Turn: On Quine's Cul-De-Sac[*]

P.M.S. Hacker
St. John's College, University of Oxford

1. NATURALISM

Contemporary American naturalism originates in the writings of Quine, the metaphysician of twentieth-century science. Like so many of Quine's doctrines, it was propounded in response to Carnap. As Quine understood matters, Carnap had been persuaded by Russell's *Our Knowledge of the External World* that it is the task of philosophy to demonstrate that such knowledge is a logical construction out of, and can be reduced to, elementary experiences. Quine rejected the reductionism of Carnap's *Logischer Aufbau*, and found the idealist basis uncongenial to his own dogmatic realist behaviourism, inspired by Watson and later reinforced by Skinner. The rejection of reductionism and an "unregenerate realism", Quine averred, were the sources of his naturalism (FME 72).

We can distinguish in Quine between three different but inter-related naturalist programmes: epistemological, ontological and philosophical. *Naturalized epistemology* is to displace traditional epistemology, transforming the investigation into "an enterprise within natural science" (NNK 68) – a psychological enterprise of investigating how the "input" of radiation, etc., impinging on nerve endings can "ultimately" result in an "output" of theoretical descriptions of the external world. I shall argue that the failure of the Russell-Carnap programme in no way implies that epistemology should be naturalized; that the project of naturalized epistemology contributes nothing to the solution of the problems

traditional epistemology struggled with; and that Quine's few forays into genuinely epistemological questions are failures.

Ontological naturalism is the doctrine that

> it is within science itself, and not in some prior philosophy, that reality is to be identified and described. (TTPT 21)

It is up to science to tell us what there is, and it offers the best theory of what exists and of how we come to know what exists. The only difference between the ontological philosopher and the scientist, according to Quine, lies in the breadth of concern: the former being concerned, for example, with the existence of material objects or classes and the latter with wombats or unicorns.

It should be noted that it is far from clear what it is to "identify and describe reality". If I identify a dandelion on the lawn, Beethoven's Opus 132 on the radio, a smell of onions in the kitchen, am I identifying "reality"? And have I done so "within science"?

In no ordinary sense of "science" is science the sole and final arbiter on what exists (e.g. Russell's childhood diaries, the pain in my leg, the Romantic movement, Mannerist style, international law, a plot to depose the king). There is no specific science that offers us the best theory of what exists, nor do the sciences collectively do so, for there is no such thing as a theory of everything that exists.

Philosophical ontology is not concerned with determining what exists in the sense in which biological taxonomy is concerned with determining and classifying what living things exist. Nor is it differentiated from a science by generality of categories. It is not as if physics is concerned to establish that mesons or quarks exist, whereas philosophy is concerned to establish that material objects or events exist. The task of ontology is to clarify, from one domain to another, what it means to say that such-and-such exists (e.g. a substance, a property, a possibility, a number, a concept, the meaning of a word, a law or legal system).

Philosophical naturalism is the view that philosophy is

> not ... an a priori propaedeutic or groundwork for science, but [is] ... continuous with science. (NNK 126)

In the USA it is widely held that with Quine's rejection of "the" analytic/synthetic distinction, the possibility of philosophical or conceptual analysis collapses, the possibility of resolving philosophical questions by a priori argument and elucidation is foreclosed, and all *good* philosophers turn out to be closet scientists. Attacks on the idea of analyticity could show that philosophy is continuous with science only if

(i)   they were successful
(ii)   philosophy consists of statements
(iii)   these contrast with scientific statements by virtue of their analyticity.

It is questionable whether Quine did successfully show that Carnap's distinction is untenable. Carnap did not think so, and explained why he did not. Grice and Strawson did not either. Quine never gave a satisfactory reply to these objections. Even in "Two Dogmas" he did not deny synonymy, and hence analyticity, in cases of stipulation, but only in the cases of ordinary terms not thus introduced. In *Roots of Reference*, he himself offered an account of analytic truths. They are those truths everyone learns merely by learning to understand them (RR 79).

Even if Quine had successfully demolished Carnap's distinction between empirical truths and truths in virtue of meaning, it would not be true that he had shown *the* analytic/synthetic distinction to be untenable, for there is not one such distinction. There is Locke's distinction between "trifling" or "barely verbal" propositions, on the one hand, and non-trifling ones, on the other, as well as Kant's, Bolzano's, Frege's and Carnap's different distinctions between analytic and synthetic truths. Their extensions are not equivalent (Kant, for example, held truths of arithmetic to be synthetic a priori, whereas Frege held them to be analytic). Some of these are epistemological distinctions, others are purely logical.

Even if someone were to demonstrate that all distinctions between analytic and synthetic propositions are untenable, it does not follow that there is no distinction between a priori and empirical propositions. Even if mathematics is not analytic, it does not follow that it is not a priori.

According to Quine,

> mathematics and logic are supported by observation only in the indirect way that those aspects [the most general and systematic] of natural science are supported by observation; namely as participating in an organized whole which, way up at its empirical edges, squares with observation. (PL 100)

But this is misconceived. Propositions of mathematics and logic are not "supported by observation". They are demonstrated by deductive proofs. It is not as if confirmation of Newtonian mechanics by celestial observations made the theorems of the calculus better supported than before. And in respect of a priority, what goes for mathematics and logic goes too for such propositions as "red is more like orange than like yellow" or "red is darker than pink". As long as we can distinguish between a tautology and a non-tautologous proposition, *and* between the specification of a measure and the statement of a measurement—the statement of a rule and the application of a rule, we can readily distinguish between what is a priori and what is empirical.

The thought that if there is no distinction between analytic and synthetic propositions, then philosophy must be "continuous" with science rests on the false supposition that what was thought to distinguish philosophical propositions from scientific ones was their analyticity. That supposition can be challenged in two ways. First, by showing that characteristic propositions that philosophers have advanced are neither analytic nor empirical (the claim of the older Wittgenstein as well as of the young Quine that there are no propositions that are true in virtue of their meanings may serve here as an example). Secondly, by denying that there are any philosophical propositions at all.

The Manifesto of the Vienna Circle, of which Carnap was both an author and signatory, pronounced that "the *essence of the new scientific world-conception* in contrast with traditional philosophy [is that] no special 'philosophic assertions' are established, assertions are merely clarified". Accordingly, the result of good philosophizing is not the production of analytic propositions peculiar to philosophy, but clarification of conceptually problematic propositions and the elimination of pseudo-propositions.

The later Wittgenstein too held that there are no philosophical propositions. The task of philosophy is to dissolve philosophical problems. These are a priori conceptual problems. They are to be tackled by the elucidation of propositions, not by their analysis into more basic ones. This requires a perspicuous representation of the problematic concepts that illuminates the problems at hand. The resultant overview does not consist of analytic propositions. This conception of conceptual analysis informed Ryle's "logical geography" of concepts and Strawson's "connective analysis", both of which were less therapeutically oriented than Wittgenstein's philosophy. None of the many philosophers who pursued conceptual analysis in this vein produced (or purported to produce) sets of analytic propositions that belong to philosophy, any more than Quine produced sets of propositions that belong to science.

Whether or not Quine's criticism of Carnap's distinction hits its target, the possibility of conceptual analysis thus understood is in no way impaired. Philosophy has not lost its proper vocation—which is not armchair science. It is categorially distinct from science, both in its methods and its results. The a priori methods of respectable philosophy are wholly distinct from the experimental and hypothetico-deductive methods of the natural sciences, and the results of philosophy logically antecede the empirical discoveries of science. They cannot licitly conflict with the truth of scientific theories—but they may, and sometimes should, demonstrate their lack of sense. *One* task of philosophy is to set straight the conceptual confusions and incoherences of scientific theories. For philosophy is neither the Queen of the sciences nor their conceptual scullery-maid, but rather a tribunal before which scientific theory may be arraigned when it trespasses beyond the bounds of sense.

2. EPISTEMOLOGY NATURALIZED

Quine ascribed to Carnap an enterprise of constructing a "first-philosophy", i.e. a form of Cartesian foundationalism, that purported to provide extra-scientific foundations for science. *Foundationalism* is the epistemological doctrine that all empirical knowledge rests ultimately on our knowledge of how things sensibly appear to us to be. Such knowledge does not itself stand in need of evidential support, but it is held to provide

the evidence for all other judgements. Carnapian foundationalism was *reductive*, i.e. it alleged that statements concerning material things are translatable into statements concerning bare experiences. The failure of the Carnapian enterprise seemed to Quine to warrant the naturalization of epistemology.

Unlike Austin, Ryle and Wittgenstein, Quine did not think that the enterprise of "bridging the gap between sense-data and bodies" was a pseudo-problem (RR 2; cf. TTPT 22). The problem was real, but the purported solution hopeless, since verification is holistic. Strict reduction and consequent eliminability of material object statements failed, according to Quine, because a "typical statement about bodies has no fund of experiential implications it can call its own. A substantial mass of theory, taken together" is required (EN 79).

So there is no need to posit sense-data to account for illusions, etc., or to posit such intermediary sensory objects of apprehension in order to account for our knowledge of material objects. The "relevance of sensory stimulation to sentences about physical objects", he declared in good behaviourist fashion, can as well (and better) be explored and explained in terms directly of the conditioning of such sentences and their parts to physical irritations of the subject's surfaces (WO 235).

Carnap's subsequent compromise of non-eliminative reduction-sentences (Ramsey-sentences) seemed to Quine pointless, renouncing the last remaining advantage of rational reconstruction over straight psychology; namely translational reduction (EN 78). "Why all this creative reconstruction, all this make-believe", he remonstrated,

> The stimulation of his sensory receptors is all the evidence anyone has to go on, ultimately, in arriving at his picture of the world. Why not just see how this construction really proceeds? Why not settle for psychology? (EN 75)

What does "settling for psychology" amount to?

First, we abandon the goal of a first philosophy prior to natural science (FME 67). Our investigation, we are told, is itself part of and continuous with natural science.

Secondly, we are called on to recognize that the sceptical challenges that epistemology has always been concerned with spring from "rudimentary science". The argument from illusion, according to Quine,

owes its force to our knowledge that sticks do not bend by immersion, and examples of mirages, after-images, dreams and the rest are, he claimed, "simply parasitic upon positive science, however primitive" (NNK 68). Consequently, in coping with these scientific problems of scepticism, we are free to use data from science and scientific knowledge (RR 3). So scientific discoveries can, without circularity or question-begging, be invoked in resolving sceptical worries.

Thirdly, epistemology thus naturalized is a branch of psychology: it studies human beings and their acquisition of knowledge or, as he put it, of "theory", investigating the relation between neural input and cognitive output (EN 83).

Hence, fourthly, naturalized epistemology, like traditional epistemology, is concerned with the relation of evidence to theory. Science, Quine averred, "tells us that our information about the world is limited to irritations of our surfaces" and the task of the scientific epistemologist is to explain how we "can have managed to arrive at science from such limited information" (FME 72).

3. EPISTEMOLOGY DENATURALIZED

Quine held Carnap's Russellian attempt to reduce our knowledge of physical objects and of other people's states of mind to the "unowned data" of elementary experience to be the culmination of traditional epistemology (FSS 13). Its failure, in his view, invited the abandonment of traditional epistemology. But no such conclusion follows. There were more variants of foundationalism than Carnap's reductivism, and *contra* Quine, there was more to traditional epistemology than foundationalism.

First, one main reason Quine gave for the failure of Carnap's enterprise was that Carnap assumed propositional as opposed to holistic verification. But in fact Carnap quite explicitly cleaved to a holistic view of theory verification and falsification, and that in a manner far closer to Duhem's modest holism than Quine's.

Secondly, it is true that Descartes, who used the Aristotelian term "first philosophy", *was* proposing a metaphysical, extra-scientific, foundation for science. The foundation he proposed involved not only our knowledge of our own thoughts (*cogitationes*) regarding how things sensibly appear to us

to be, but also truths of reason known by the natural light, knowledge of simple natures and a proof of the existence of God. But Descartes's foundationalism was in no sense reductive, and the failure of Carnapian reductivism is irrelevant to Cartesian foundationalism. Lockean foundationalism is different again, and is akin to inference from the data of sense, i.e. ideas, to the best explanation for such data. This too was not reductive, and its latter-day heirs (e.g. J. L. Mackie's account) are untouched by the failure of Carnapian reductivism. So the failure of Carnapian reductivist foundationalism in itself does not even imply the bankruptcy of other foundationalist enterprises, let alone the abandonment of traditional epistemology.

What *was* wrong with Cartesian and Lockean foundationalism was not reductivism (since they were not reductive), but the foundationalist base. This objection applies equally to Carnapian reductivism. The thought that the foundations of our knowledge of the external world lie in our knowledge of our own subjective experience, in how things subjectively seem to us to be or in the ideas with which the mind is furnished by experience, is misconceived. For the attempted philosophical justifications of "our knowledge of the external world" in the foundationalist tradition involved radical misuses of a wide range of verbs of sensation, perception and observation, and their manifold cognates. Foundationalism presupposes the intelligibility of a logically private language. Moreover, it misconstrues the actual role of sentences of the form "It seems to me just as if *p*" or "It appears to be an M" and of the sentence-forming operators "It seems that ...", "It appears to be ..." and "It looks as if ...". Finally, the reductive base presupposes objective spatio-temporal reference and simultaneously makes it impossible. Foundationalism (reductive and non-reductive alike) is not, as Quine asserted, an intelligible failure for holistic reasons, it is an unintelligible endeavour rooted in Cartesian misconceptions about knowledge, doubt and certainty, and in mistaken Cartesian strategies of combating scepticism on ground of its own choosing—namely the quest for certainty.

So, foundationalism *is* to be rejected. But why should the naturalization of epistemology follow? The only reasons Quine gave are inadequate.

(1) Admitting that naturalized epistemology is "a far cry from old epistemology", he held that it is an "enlightened persistence" in the

original problem (RR 3). The original problem was: how can we justify our claims to know anything extra-mental? The allegedly enlightened transform is: how does it come about that we know anything extra-somatic? That question, Quine held, is a question for psychology, which will explain how sundry irritations of our surfaces ultimately result in true statements of science. Naturalized epistemology will be concerned with elaborating *causal links* between the "input" of sensory stimuli and the output of statements describing the external world. The proper task of scientific epistemology must perforce be allocated to future neuropsychology.

It is mistaken to suppose that there is anything enlightened about substituting a causal question about the ontogeny of human knowledge for conceptual questions concerning the general categories of knowledge and the kind of warrant or justification that non-evident beliefs may require. The question of what *warrants* a claim to knowledge concerning objective particulars is not resolved by an explanation of what are the causal processes necessary for attaining any such knowledge. Indeed, the causal investigation *presupposes* that sceptical qualms can be laid to rest, but are no substitute for laying them to rest.

The sceptical qualms that, in Quine's view, are the source of traditional epistemology, arise, according to him, from "science" (empirical knowledge), and in answering them, he claims, we are free to appeal to scientifically established fact (agreed empirical knowledge) without circularity (RR 3). That is mistaken. What we have to do is to show that the sceptic's arguments and presuppositions are awry.

Quine rarely ventured into the territory of epistemological scepticism, but when he did, his forays lacked penetration. To scepticism about dreaming, he responded: "I am ruling the dream hypothesis out in the sense that I dismiss it as very unlikely". To the updated variant of dream-scepticism that one may be a brain in a vat, Quine responded:

> I would think in terms of naturalistic plausibility. What we know, or what we believe ... is that it would really be an implausible achievement, at this stage anyway, to rig up such a brain. And so I don't think I am one. (Fogelin 2004, 43f.)

I don't think that Quine quite understood the point. Scepticism is not a challenge to one of the planks in Neurath's boat. It is a challenge to the logical possibility of seafaring. And it cannot be answered by invoking "scientific" facts or common sense, or by pointing out that boats do actually go to sea. (One cannot resolve Zeno's paradox by observing that Achilles can overtake the tortoise by putting one foot down after another.) The problems it raises are purely conceptual ones, and they are to be answered by purely conceptual means—by clarification of the relevant elements of our conceptual scheme. This will show what is awry with the sceptical challenge itself.

(2) The second reason Quine gave for opting for naturalized epistemology is that

> If all we hope for is a reconstruction that links science to experience in explicit ways short of translation, then it would seem more sensible to settle for psychology. Better to discover how science is in fact developed and learned than to fabricate a fictitious structure to a similar effect. (EN 78)

But the failure of Carnapian reductive foundationalism has no such implication. If the reductive enterprise fails, the first thing called for is a *philosophical* investigation into the reasons for the foundationalist project in the first place. This may reveal that the questions were based on misconceptions. Quine held that the question of whether there is an external world is a bad question. But, like Hume, he claimed that the question that replaces it is "whence the strength of our notion that there is an external world?" (SLS 217). In his view, the existence of external objects in the physical world is an efficient posit. "In a contest for sheer systematic utility for science", he wrote, "the notion of physical object still leads the field" (WO 238). The epistemological enterprise of trying to justify our knowledge of the external world in the face of sceptical challenges is to be replaced by a scientific explanation of the causal processes that lead to our positing objects and acquiring our "theory of the world". That is mistaken: we do not "posit" objects, and we do not have a "theory of the world".

It *is* correct that foundationalism in its various forms, is misconceived. But it is incorrect to suppose that once it is rejected, there is nothing left for epistemology to do than become scientifically naturalized. There is a great

deal more to epistemology than answering the sceptic. Contrary to what Quine asserted, what prompted epistemology was not to see how evidence relates to theory. It was, above all, to explain what knowledge is, what its characteristic marks are and what difference there is between knowledge and opinion. It was to investigate the scope and limits of knowledge; to determine whether humanity can achieve any absolute knowledge or whether all knowledge is relative; to discover whether pure reason alone can attain any knowledge of the world; to decide whether absolute certainty is obtainable in any of the forms of knowledge attainable by us; to show whether moral knowledge is attainable, whether mathematical knowledge is more certain than perceptual knowledge, whether we can know that God exists or whether the soul is immortal. And so on.

Early epistemology focused on the different sources of knowledge and on the different kinds of knowledge that we can attain. Despite Quine's avowals to the contrary, there are radical differences between mathematical knowledge and empirical knowledge, between self-knowledge and knowledge of others, between knowledge of objects and knowledge of scientific theory (e.g. of electricity, magnetism, ionic theory), between the natural and the social sciences, and so forth. It would be a mistake to suppose that one can glibly say, knowledge is knowledge —it merely has different objects. Knowledge that Jack is taller than Jill is *categorially* unlike knowledge that red is darker than pink. To know the difference between right and wrong is *radically* unlike knowing the difference between Coxes and Bramleys. To know what I want is *epistemologically* unlike knowing what you want, and to know what I think about a given question is not akin to knowing what you think. Could naturalized epistemology contribute to the clarification of such conceptual differences? I think not—any more than mathematics naturalized could explain the differences between natural numbers and signed integers, or between rationals and irrationals.

Traditional epistemologists want to know whether knowledge is true belief and a further condition (as was supposed in mid-twentieth century), or whether knowledge does not even imply belief (as was previously held). We want to know when knowledge does and when it does not require justification. We need to be clear what is ascribed to a person when it is said that he knows something. Is it a distinctive mental state, an

achievement, a performance, a disposition or an ability? Could knowing or believing that *p* be identical with a state of the brain? Why can one say "he believes that *p*, but it is not the case that *p*", whereas one cannot say "I believe that *p*, but it is not the case that *p*"? Why are there ways, methods and means of achieving, attaining or receiving knowledge, but not belief (as opposed to faith)? Why can one know, but not believe who, what, which, when, whether and how? Why can one believe, but not know, wholeheartedly, passionately, hesitantly, foolishly, thoughtlessly, fanatically, dogmatically or reasonably? Why can one know, but not believe, something perfectly well, thoroughly or in detail? And so on— through many hundreds of similar questions pertaining not only to knowledge and belief, but also to doubt, certainty, remembering, forgetting, observing, noticing, recognising, attending, being aware of, being conscious of, not to mention the numerous verbs of perception and their cognates. What needs to be clarified if these questions are to be answered is *the web of our epistemic concepts*, the ways in which the various concepts hang together, the various forms of their compatibilities and incompatibilities, their point and purpose, their presuppositions and different forms of context dependency. To this venerable exercise in connective analysis, scientific knowledge, psychology, neuroscience and self-styled cognitive science can contribute nothing whatsoever.

Quine rarely paid attention to such questions. But when he did his answers were *not* essays in naturalized epistemology, i.e. parts of empirically testable theories, but patently traditional philosophical claims. They were, equally patently, inadequate. I shall give three examples.

"Knowledge", Quine wrote, "connotes certainty" (Q 109), and rightly hesitated before limiting knowledge to the absolutely certain. But knowledge does not connote certainty at all. Rather, it is improper *to claim to know* something if one has *doubts*. A legitimate claim to knowledge presupposes absence of doubt (not presence of certainty), but knowledge as such does not (we do not fail doctoral students in their oral examinations because of their uncertainty).

Faced with the Gettier counter-examples to the definition of "knowledge" as justified true belief, Quine did not even try to show how they can be accommodated within an alternative account of knowledge, but rather concluded:

> I think that for scientific or philosophical purposes the best we can do is give up
> the notion of knowledge as a bad job and make do with its separate ingredients.
> We can still speak of belief as being true, and of one belief as firmer or more
> certain, to the believer's mind, than another. (Q 109)

One wonders what philosophical or scientific purposes Quine had in mind. In truth the concept of knowledge is not an isolated dangler in our epistemic conceptual scheme that can be excised without collateral damage. Did Quine also want to give up the notion of memory (knowledge retained) as a bad job? Are neuroscientists investigating clinical aphasic syndromes following lesions to Wernicke's and Broca's areas in the cortex not investigating the neural foundations of memory? Did Quine also wish to give up the notions of perceiving that *p* (in its various forms), being aware, being conscious, recognizing, noticing that *p*—all of which imply knowing that *p*? These cognitive concepts too are integral to cognitive neuroscience and experimental psychology.

If we are to give up the notion of knowing, at least we retain that of believing. What, according to Quine, is that? "Belief", he claimed, "is a disposition" (Q 18). The dispositions of which he holds the mind to consist "are dispositions to behave, and those are physiological states". Hence he ended up, he said, "with the so-called identity theory of the mind: mental states are states of the body" (MVD 94). But this too is mistaken. Beliefs (i.e. believings) are not dispositions to behave. Dispositions are essentially characterized by what they are dispositions to do, beliefs are essentially characterized by reference to what is believed to be so. To explain human voluntary behaviour by reference to a person's dispositions is to explain it by reference to his nature, temperament or personal traits. To explain A's voluntary V-ing by reference to his belief that *p* is not to explain it by reference to his traits of character; but nor is it to explain it by reference to his behavioural habits, tendencies or pronenesses (which is what Quine meant by "disposition"). It is to explain it in terms of what A took as his reason for V-ing. To know that A has a certain disposition (in Quine's sense) is to know that he is prone or liable to act or react in certain ways in response to certain circumstances. But one can know that A believes that *p* without knowing what, if anything, A is prone or liable to do. The utterance "I believe that *p* but it is not the case that *p*" is a kind of

contradiction. But "I have a disposition (I tend, am inclined or prone) to V, but it is not the case that *p*" is not a contradiction of any kind. If A believes that *p*, then it follows that A is right if *p* and wrong if not-*p*, but no such thing follows from A's having a behavioural disposition, tendency or proneness.

Quine compounds his errors by identifying a disposition with its vehicle, claiming that the human dispositions are physiological states of the body or brain. But a disposition, no matter whether an inanimate one or a human one, is never identical with its vehicle, any more than an ability is identical with the structures that make it possible (Kenny 1975, 10f. and Kenny 1989, 72f.). The horsepower of the car is not beneath its bonnet, and the intoxicative power of whisky is neither lighter nor heavier than the constituent alcohol that is its vehicle. So even if it were true that believing that *p* is a disposition, proneness or tendency, it would not follow that it is identical with a neural state. For were believing that *p* identical with a neural state, one would be able to say "I believe that *p* (referring thus to one's neural state), but it is not the case that *p*".

In short, the alternative to Carnapian reductionism is not naturalized epistemology. Naturalized epistemology does not answer the great questions of epistemology and is no substitute for their answers.

4. REFERENCES AND ABBREVIATIONS

Fogelin, R. 2004: "Aspects of Quine's Naturalized Epistemology." In: Gibson, R. F. (ed.) *The Cambridge Conpanion to Quine*. Cambridge: Cambridge University Press.

Kenny, A. J. P. 1975: *Will, Freedom and Power*. Oxford: Blackwell.

Kenny, A. J. P. 1989: *The Metaphysics of Mind*. Oxford: Clarendon Press.

Quine, W. V. O. 1960: *Word and Object*. Cambridge, MA: MIT Press.

Quine, W. V. O. 1966: "Mr Strawson on Logical Theory." Repr. in Quine, W. V. O.: *The Ways of Paradox and Other Essays*. New York: Random House.

Quine, W. V. O. 1969: *Ontological Relativity and Other Essays*. New York: Columbia University Press.

Quine, W. V. O. 1970: *Philosophy of Logic*. Englewood Cliffs, N.J.: Prentice-Hall.

Quine, W. V. O 1974: *Roots of Reference*. La Salle, IL: Open Court.

Quine, W. V. O. 1975: "The Nature of Natural Knowledge." In: Guttenplan, S. (ed.) *Mind and Language*. Oxford: Clarendon Press, 67-81.

Quine, W. V. O. 1975: "Mind and Verbal Dispositions." In: Guttenplan, S. (ed.) *Mind and Language*. Oxford: Clarendon Press, 83-95.

Quine, W. V. O. 1981: "Five Milestones of Empiricism." Repr. in Quine, W. V. O.: *Theories and Things*. Cambridge, MA: Harvard University Press, [SEITENZAHLEN].

Quine, W. V. O. 1981: "Things and their Place in Theories." Repr. in Quine, W. V. O.: *Theories and Things*. Cambridge, MA: Harvard University Press, [SEITENZAHLEN].

Quine, W. V. O. 1990: *Quiddities: An Intermittently Philosophical Dictionary*. Penguin Books, London.

Quine, W. V. O. 1995: *From Stimulus to Science*. Cambridge, MA: Harvard University Press.

EN – 'Epistemology Naturalized', in *Ontological Relativity and Other Essays*.

NNK – 'The Nature of Natural Knowledge.' In S. Guttenplan (ed.) *Mind and Language*.

MVD – 'Mind and Verbal Dispositions.' In S. Guttenplan (ed.) *Mind and Language*.

FME – 'Five Milestones of Empiricism.' Repr. in *Theories and Things*.

FSS – *From Stimulus to Science*.

PL – *Philosophy of Logic*.

Q – *Quiddities: An Intermittently Philosophical Dictionary*.

RR – *Roots of Reference*.

SLS – 'Mr Strawson on Logical Theory.' In *The Ways of Paradox and Other Essays*.

TTPT – 'Things and their Place in Theories' In *Theories and Things*.

WO – *Word and Object*.

The Heavy Burden of Proof for Ontological Naturalism

Georg Gasser & Matthias Stefan
University of Innsbruck

1. WHY NATURALISTS CANNOT DODGE ONTOLOGY

Naturalism enjoys almost a status of orthodoxy among contemporary analytic philosophers. Unfortunately naturalism is not a clearly defined philosophical position. Rather it is comparable with a broad stream consisting of various philosophical approaches. It is not our aim to provide an elaborated definition of contemporary naturalism. Following Mike Rea's characterization (Rea 2002, 50-73) we consider naturalism to be a programmatic set of strategies to understand the world. Central to this programmatic set of strategies is to analyse and present the world by relying heavily on science (Forrest 1996, 89). Hence, naturalism implies the attitude to consider science (at least) as the primary source of reliable knowledge about reality. Science, according to naturalism, has shown to be the most successful strategy for understanding the structure of our world and its causal interaction (Löffler 1999, 36). This is the way we understand naturalism. Such a characterization is very vague. It leaves room for many different interpretations methodological, epistemological and ontological alike. In our paper we focus on ontological issues. We aim at exploring which ontological commitments come hand in hand with a naturalistic outlook on reality. This exploration presupposes that naturalism and ontological reflection are intertwined. An argument supporting this presupposition goes as follows:

Ontology is the philosophical discipline investigating the ultimate structures of reality. It aims at formulating what exists and what should be deemed as real or unreal. If it is true that naturalists claim that the methods of natural science exert a kind of hegemony over all strategies pursuing

truth, then sciences are best in telling us what exists. Consequently ontology—according to naturalism—depends on what is recognised as real by sciences. A naturalistic minded philosopher, if asked what the ultimate structures of our world are, should consistently base his/her answer on current scientific research and the ontological commitments coming along with it.[1]

Many contemporary naturalists, however, refuse to pursue studies in ontology. They argue that naturalism should be interpreted neutrally from any ontological point of view. Ontological discussions tend to complicate—instead of enhancing—the interdisciplinary project between science, philosophy and common sense respectively (e.g. Tetens 2000, 287f.; Clayton 2004, 142). If naturalism leaves ontological discussions aside and concentrates on epistemology and methodology, for instance, it might support interdisciplinary projects between science and philosophy. It could start with concepts of folk psychology and take into account empirical investigations of cognitive science. In such a way naturalism might be helpful to analyse and better understand central concepts of cognitive science (Koppelberg 2000).

Whatever the specific merits of epistemological or methodological naturalism might be, they avoid the philosophical puzzles arising from ontological thought. Such or similar the argument goes against the view of taking ontological issues seriously within naturalism.

We concede that most scientists do not explicitly care for ontology. It is alien to scientific practice to spell out what a certain theory implies ontologically. This is rightly done. Scientists are not paid for doing ontology. However, we deem an ontologically neutral or abstinent naturalism unsatisfying from a philosophical point of view. If one accepts scientific realism and the thesis that explanatory concepts in science come along with ontological commitments, then a tension between different causal claims and various scientific explanations becomes a problem at some point. Psychology, for instance, makes causal claims about mental states and neuroscience makes causal claims about neurological facts.

---

[1] Papineau 2001, for instance, aims at pointing out how modern physics led many philosophers to become ontological physicalists. Dupré 2004, 37ff., portrays a similar route from materialism over modern physics to physicalism.

Now, if it is assumed (as most non-dualists would do) that mentality depends upon neuronal activity in the brain, then the question arises concerning the relation of these different causal claims. Are they just two different descriptions of the same reality? The decisive question is who or what is doing the causal work: Does mentality dispose of mental causal powers? Or is mentality only causally efficacious in virtue of its dependence upon neurological activity? This problem shows that it is not unusual that various scientific explanations and its corresponding ontological implications are in conflict with each other. There are causal entanglements between the various levels of reality investigated by different scientific disciplines. These causal entanglements raise all sorts of issues which have been dubbed "problem of causal competition", "problem of causal exclusion" or "problem of cross-level causation". These issues will have to be evaded or answered. Because it has proven to be difficult to provide a convincing way to evade the tensions between different causal stories at the various levels of reality, we assume that an ontological neutral position can only be provisional for naturalism: It might be useful to leave ontological implications out of account for specific aims of scientific research or philosophical reflection. However, if it is believed that scientific concepts and hypotheses refer to something real, then methodological and epistemological issues are closely intertwined with ontological assumptions. Ontological questions are neither external to scientific practice nor of no interest for science. Scientific theories make statements about entities being causally efficacious. By doing so, ontological questions are implicitly raised by scientific theories. If our argument is correct, naturalistic minded philosophers should regard it as a substantial topic of their research to work out which ultimate structures of reality we are reasonably enabled to accept according to naturalism. Ontology is not a marginal but a central issue for naturalism.

2. NATURALISTS' COMPLAINT AND MEN OF STRAW

After having argued for the importance of taking ontology seriously in naturalistic thought, in this section we discuss briefly which ontological options should be preferred over others. In surveying philosophical literature on naturalism we encounter again and again a complaint from the

side of naturalists (see for instance Sukopp 2006, 28-33). Many naturalists complain they are attacked from two sides: First, they are attacked because naturalism is presented as being ultimately radically reductionist or eliminativist. Secondly, they are attacked because naturalism is presented in such a liberal way that almost everyone is willing to accept it.

According to the first line of attack naturalism coincides with reductive physicalism. Based upon contemporary physical theories it is claimed that our world contains only basic micro-physical entities. These entities together with fundamental physical laws are sufficient to account for all phenomena we encounter in our universe. Once the physical facts at the bottom level of reality and the laws holding them are fixed, all facts of our universe are fixed. There is nothing relevant above these fundamental physical facts. Such a position implies the (in principle) reduction of our macro-world to micro-physics, the (in principle) reduction of laws of special sciences to the fundamental laws of physics and the (in principle) reduction of 'non-physical phenomena' such as consciousness, subjectivity or intentionality to physical phenomena. In such a reductionist world ultimately only micro-physical entities and their causal interactions exist. For many people such an ontology has unpalatable consequences because the world we are familiar with is ontologically inferior or negligible. There are no physical facts above micro-physics, no causal powers above the powers of the ultimate constituent parts of reality. The world as we conceive it, the causal powers we ascribe to the objects of our meso-cosmos, in short, our Lebenswelt, drains away after all. In the final analysis the ultimate level of reality—the level of micro-physics—is the only level which truly has to be taken seriously from an ontological (and maybe on the long run also from a scientific) point of view.

According to the second line of attack, ontological naturalism is so widely defined that nearly every ontological thesis is part of it. If only God, angels, immaterial substances or mythical creatures are excluded from a naturalistic ontology, then naturalism does not seem to offer any interesting ontological insights. Anyone who has not strong theistic, animistic or obscurantist tendencies would subscribe to such a version of ontological naturalism. A naturalism telling us that our world is material based, that it evolved over time in a continuous process, that complex systems consist of simpler parts, and that we do not have to postulate non-

material entities for explaining our universe, is a rather unimpressive philosophical program. It might legitimately be asked whether such a view is still worth to be named naturalism (Stroud 2004, 34f., Keil and Schnädelbach 2000, 9f.).

Most naturalists consider the two options sketched above as mere men of straw which easily can be torn into pieces. This complaint from the naturalist's side implies that any serious analysis (and critique) of naturalism should look for positions being situated between a too strong reductionist physicalism and a too liberal naturalism. We take these complaints seriously. In the next section we aim at spelling out possible compromises between the two extremes.

3. SCIENTISM AND ITS ONTOLOGICAL OPTIONS

We presented naturalism as being committed to the attitude that science ultimately recognizes what is real and unreal. Any ontology faithful to this attitude has to be developed under the authoritative guidance of science. We label such an attitude 'scientism'. Kornblith gives a succinct expression of this view:

> Current scientific theories are rich in their metaphysical implications. The task of the naturalist metaphysician, as I see it, is *simply to draw out the metaphysical implications of contemporary science*. A metaphysics which goes beyond the commitments of science is simply unsupported by the best available evidence. (Kornblith 1994, 40, our italics).

Kornblith's view is programmatic in character. Hence, let us ask: How shall an ontology be construed out of contemporary scientific theories? We suggest that first of all we have to clarify *which sciences* are to be accepted as providing relevant information for a naturalistic ontology. Second, we have to explain how the relevant sciences *relate* to each other. It is easy to see that these two problems are connected. A well-known example from the history of science helps to explicate this interconnection. The bonding problem in chemistry was a much debated topic at the edge of the 19th century. As long as micro-explanations of chemical bonding were not at hand, chemical theories assumed fundamental chemical forces of chemical elements. After the development of quantum mechanics, the gap between

chemistry and physics began to shrink. It was detected that quantum states and not some emergent chemical properties are the reason for chemical forces. Chemical bonding became explainable in terms of quantum mechanics.[2] The important point for our discussion is that in this case all relevant information is provided by physical theory. Physics tells us everything about the problem of chemical bonding. This scientific progress had influential consequences for ontology. Before quantum states were known, emergent chemical properties were assumed as being part of our world. After the discovery of quantum states the assumption of proper chemical properties became superfluous. As the example shows, it might be the case that within the disciplines, scientism considers as relevant for its ontological program, specific theories of a determinate discipline (e.g. chemistry) are reducible to more basic theories in another (e.g. physics). For a naturalist subscribing to scientism it is not only important to identify those sciences which provide relevant information for a naturalistic ontology; it is crucial to elucidate their interrelationship as well. We propose three solutions for achieving this aim:

(i)   Scientism pursues a reductionist strategy. Scientism assumes that the entities of higher level sciences are reducible to micro-physical entities. Biological entities, for instance, ought to be reduced to chemical entities and these to physical ones. Scientism then turns into *physicalism*, as all higher level sciences are nothing more than special cases of (an assumed complete) science of physics.

(ii) Scientism becomes a kind of *conciliable naturalism*. Conciliable naturalism says we should accept everything as relevant that we think we need to make sense of and which we are convinced is part of our world (Stroud 2004, 33). Conciliable naturalism accepts the whole range from natural to social sciences, and humanities.

(iii) Scientism relies on some well established sciences, such as physics, chemistry and biology. We call this position

---

[2] McLaughlin 1992 discusses these scientific discoveries at length and relates them directly to the rise and fall of British Emergentism.

'naturalism of core sciences'. It might be claimed that this version of naturalism is some kind of a compromise: It avoids the openness of conciliable naturalism on the one hand without subscribing to a too strong reductionist version of physicalism on the other hand.

How shall we deal with the three options at hand? Position (i) is often considered problematic. It seems to coincide with the position identified above as our first man of straw because it claims that (micro-)physics alone provides the relevant information for any ontology. Position (ii) seems to be a non-starter due to its open-mindedness. Conciliable naturalism is not more than "a slogan on a banner raised to attract the admiration of those who agree that no supernatural agents are at work in the world." (Stroud 2004, 35) Conciliable naturalism is identifiable with the position identified above as our second man of straw.

The remaining candidate is position (iii). For our argument it is not of further importance whether 'naturalism of the core sciences' considers only physics, chemistry and biology as relevant sciences or whether the list can be extended[3]. Central for our argument is that a well-defined notion of sciences seems to be presupposed. Explicating the concept of science is a necessary precondition for being able to say which sciences take part in the ontological undertaking of naturalism. It is, however, anything but clear what natural sciences are. As long as this problem remains unsolved the problem of sciences' interdependency cannot be tackled either. Without a clear concept of science, scientism can hardly justify why it takes certain sciences seriously for ontology, whereas others are seen as less important. This gives rise to the impression that an envisaged science based ontology amounts to a mere matter of taste—the groundless capriciousness of certain people to favour certain disciplines over others.

---

[3] Of course, the list cannot be extended arbitrarily. This extension has to be stopped at some point before turning this position into a form of conciliable naturalism. But this problem is of no further importance here.

Naturalists have to answer the question about their concept of science:

> Having declared that the methods of natural science provide the only avenue to truth, the naturalist should be prepared to say what these methods are, or which sciences qualify as 'natural sciences.' (Keil 2000, 148)

There are three possibilities to develop such a concept of science:

(a) Providing methodological criteria which separate sciences from non-sciences.
(b) Providing a list of acceptable sciences.
(c) Demonstrating the unity of science.

Let us discuss the three possibilities in turn.

(a) It is a fact that in scientific progress methods of science develop. Long established methods in scientific practice are factored out as not being scientific enough anymore whereas other methods become integral parts of current scientific practice. Standards what counts as scientific and what as unscientific change with the course of science's development. In short, the methods of science cannot be determined a priori. Our characterization of naturalism implies that naturalism finds its orientation within science and in this respect it is at science's mercy. As a consequence, naturalism cannot impose a priori methodological criteria on science. Otherwise naturalism is not a loyal companion of science anymore but sets itself up as judge over it. Such a move is inconsistent with naturalism's commitment to follow and cooperate with science. If this argument is sound, naturalism cannot develop criteria for distinguishing sciences from non-sciences because such criteria would determine a priori what has to count as science and what not.

(b) Providing a list of admissible sciences has also to be refused. It is impossible to provide a non-arbitrary list of accepted sciences without methodological criteria. If we look for methodological criteria, we are back at (a), which has been already ruled out as a possible solution to the problem at hand.

(c) Demonstrating the unity of science seems to be a philosophical project most evidence speaks against. We have little reason to believe in any kind of unity of science. If we look at science as practiced, then we are

unable to notice any methods of investigation which are characteristic for all sciences. There are also naturalistic attempts to provide a unity of sciences in content, which assumes that all sciences investigate entities that are one and the same in nature. These entities might be, for instance, physical in nature, so that all sciences are united in having physical content. Such a move brings us, finally, to the doctrine of physicalism which we have discussed as a first possible interpretation of scientism. Whether this is a viable way we have to let others decide. Currently such attempts lack any convincing philosophical basis and empirical support. There are no signs of a realisation of the project of the unity of science in terms of method and content (Dupré 2004, 51).

Scientism seems to be incapable of giving an adequate concept of science. Without such a concept, however, it is impossible to construe an ontology out of contemporary scientific theories. And even if we assume, for the sake of argument, that scientism possesses such a concept of science, an abiding problem for a science-based ontology still remains, namely that of the interdependency of sciences. Most sciences imply 'local ontologies'. A local ontology is the ontology a scientific discipline (or even a specific theory within a discipline) uses explicitly or implicitly for its area of research. Biology, for instance, (to speak simplified) works among other ontological categories with three-dimensional objects, such as organisms for explaining biological phenomena. Many of these objects of biological research correspond to objects we are familiar with from common sense. Particle physics instead might carry out its research in a four dimensional time-space system with fields, atoms and electron clouds. It might not feel the urge to refer to three-dimensional objects familiar to us from everyday experience.[4] It is of no further importance for our argumentation what entities exactly are assumed in different scientific areas. It suffices to point to the fact that different local ontologies in different sciences lead irreversibly to an unpleasant consequence: Various sciences use different ontological categories, while their interrelationship is everything but clear. Advocates of scientism have to explain how different

---

[4] It should be noted that this assumed ontological framework for physics, which is already simplified, is not at all undisputed. Within physics itself are many sites of (epistemological and ontological) fracture (see Falkenburg 2006).

ontological categories, drawn out of contemporary science, relate to each other. If a clarification of these relationships cannot be provided, the project of scientism is seriously threatened. It is as if we had many pieces of a puzzle and did not know how to put them together to a unified picture.

A naturalist might reply that a thoroughgoing science-based ontology is not available. There are gaps in a science-based ontology. As the various sciences are not united so the local ontologies are not related to them. These gaps have to be accepted as expressions of our ignorance within the project of a science-based ontology. All we have are small puzzles, but a great unifying picture is simply beyond our reach. We ought to live with a fragmented ontology of our reality.

We see the main problem of such a proposal in the acceptance of gaps. As argued before, scientism needs to draw a line between acceptable and unacceptable sciences for being able to construe its ontology. Often this line is drawn between natural sciences and other academic disciplines which have the mental or the social as their primary research object (Mellor/Crane 1995). Let us suppose there are reasonable grounds to draw the distinction at the intersection of those sciences concerned with the physical and those studying the assumed non-physical. If, for the sake of argument, we accept this distinction we should be prepared to answer the question why we are willing to tolerate gaps among the natural sciences and still adhere to the drawn distinction. Why should we consider a fragmentation between particle physics, atomic and molecular theory, biology, physiology, or neurology as less problematic than the gap between the physical and the mental? This assumption grounds on the presupposition of the unity of the accepted sciences. It is presupposed that one can smoothly go up the hierarchy of sciences from physics over chemistry and biology to neurobiology without any change in content. It is always the same realm that is investigated. The only problematic gap, then, is lurking between the physical and the mental. Implicitly it is presupposed that the mental stands alone in our physical world (Churchland 1981, 75). This presupposition, however, is by itself not justified, as the unity of science has to be provided first. And we saw already that the signs of a realisation of the project of the unity of science are currently few and far between.

If scientism goes another route instead and accepts gaps for its ontological undertaking within sciences itself, as well as a further gap between so called hard and soft sciences, then a science based ontology does not have any reason to favour the subject matter of hard sciences over the one of soft sciences. Scientism, then, would accept a plurality of sciences as equally relevant for ontology. There would not be one ontology but many ontologies depending on the respective (scientific) perspective one takes up for getting an accurate account of the phenomena under investigation. It seems likely to us that such a view turns into liberal naturalism which has been ruled out as a non-starter at the very beginning of the discussion. What our discussion should have made clear is the following:

(i) If almost everything is assumed as being part of nature, then naturalism becomes so liberal that it turns into triviality.

(ii) There is no generally accepted concept of science which allows drawing a clear line between acceptable and non-acceptable sciences. Criteria to distinguish between relevant and irrelevant sciences seem to be difficult to obtain. Without such criteria, however, the entire project of a science-based ontology cannot be achieved. It remains unclear to which sciences ontology should refer.

(iii) Even if it could plausibly be argued which sciences a naturalistic minded philosopher should take into consideration for his/her ontological studies, the interrelationship between the accepted sciences and their assumed entities has still to be clarified.

Drawing out the ontological implications of contemporary science, as Kornblith demanded, has shown to be full of flaws. Scientism as a philosophical project seems to be a failure—at least from the perspective of its ontological implications. How shall we proceed then? We suggest returning to one option we presented at the beginning of our discussion: physicalism. It seems to be the only remaining option providing a solution for the problems of the notion of science and the interdependency between scientific disciplines. If this were true, naturalists should accept what they

are generally loath to do: To argue that physicalism, and the reductionism coming along with it, is the most promising route for presenting a distinctive naturalistic ontology.

4. ONTOLOGICAL PHYSICALISM

How shall physicalism be characterized? We start with a list of core tenets physicalism[5] entails (see e.g. Pettit 1993, 213-223; Beckermann 2000, 128-143; Kim 2005, 149f.):

(1)   The world is constituted out of microphysical entities which physics is in the best position to identify. Every entity in the world is either a microphysical entity itself or constitutes of microphysical entities.

(2)   Microphysical entities are subject to law-like regularities described by physics. Since microphysical entities constitute everything, macro-entities are subject to laws which are constituted by microphysical laws.

(3)   Once microphysical entities and law-like regularities holding them are fixed, all facts are fixed—metaphysically speaking (Loewer 2001a, 39). This is implied by (i) and (ii).

(4)   Higher level entities exist by being identical to or by *supervening upon physical entities.* This thesis itself remains tacit whether we can still be realistic about higher level entities (Kornblith 1994, 42; Loewer 2001a, 46; Hüttemann/Papineau 2005, 34).

---

[5]   Physicalism is an ambiguous term. Some philosophers call themselves physicalists but in fact they reject only the acceptance of non-material substances in our world. Such a version of physicalism is identifiable with what we call 'liberal or conciliable naturalism'. We think it is ill-founded to label such a view 'physicalism' as it creates more confusion than clarifications. Others, however, propose some constrained notion of physicalism, which we aim at defining in points (1) to (6). We leave it open which commitments the single philosophical tenets entail. The point is rather, that everyone accepting tenets (1) to (6) faces the problem we expose in what fallows.

(5) Causation of non-fundamental entities exists *supervening* upon *or depending on* physical facts and laws. This thesis does not rule out that higher level causation exists as well.

(6) Physicalim's commitment to the thesis that higher level entities are *sums of* or *supervenient upon* physical entities leaves open whether higher level entities are only adequately studied and interpreted via a reductionist methodology and ontology.

Tenets (1) to (6) leave room for a variety of physicalistic positions. A certain prevalence for microphysical entities and the objects studied by physics can be noticed as general commitment of them all. But at this point physicalism is still open to various interpretations reductionist and non-reductionist alike.

Reductionists claim that all higher level entities are (in principle) reducible to physical ones. According to reductionism, higher level entities present no domain of their own but are reducible to and identical with entities in the physical realm. In a final analysis, all that exists are physical entities—whatever they may be—and sums of them.

Non-reductive physicalists reject a strong reading of (1) to (6). They accept the existence of higher level entities in a genuine sense as well.[6] According to non-reductive physicalists there are facts in the world that simply cannot be stated or noticed in terms of lower level entities. It is the failure of reduction because of the incompleteness of lower level ontology that justifies the acceptance of irreducible higher level entities. These higher level entities are asymmetrically dependent on the physical level. This dependency-relation is mostly dubbed as supervenience relation. Basically it says: No changes at the higher levels without changes at the lower level. Two systems exemplifying exactly the same physical states exemplify the same higher level states as well, but not the other way around.

What should be noted at this point is that reductionists and non-reductionists share a common worry: To leave out important features of our world which we care about. If reductionists aim at reducing higher

---

[6] For a detailed version of non-reductive physicalism see e.g. Poland 1994.

level entities to lower ones then they do so for being realists about higher level entities. By providing a reduction, higher level entities no longer 'hover over' the physical world but become a part of it. The existence and causal efficacy of higher level entities is thus guaranteed by reducing them. The worry of most reductionists is that entities which cannot be reduced will prove to be epiphenomenal or, even worse, unreal altogether (Kim 1995). Exactly the same worry drives non-reductionists as well. They hold, however, the opposite view of what counts as legitimizing higher level entities. For non-reductionists reduction amounts to elimination of higher level entities: If mental states are reducible to physical states, then there are only physical states. Mental states become superfluous, ontologically speaking. To protect the ontological status of higher level entities it has to be shown that they are indispensable and irreducible. Any ontology leaving them out or not granting them the full right to exist would be incomplete: Important parts of our world would be missing.

The reason we stress this common worry of reductive and non-reductive physicalists alike is to avoid a 'straw man argument' in the discussion. Often anti-naturalists argue that naturalists do not ascribe great importance to such crucial features of our self-conception as subjectivity or the first-person-perspective. Such a reproach misses the mark. Many naturalists aim at naturalizing these features via reductionist strategies. Only few consider them as entirely eliminable or superfluous. Thus, reduction should not be confused with elimination (see Kim 2005, 160). Naturalization via reduction means to provide a home for ontologically disputable entities such as mental entities within an ontologically undisputed realm such as the physical.

Probably most philosophers fancying physicalism, subscribe to a version of non-reductive physicalism. According to them one can go up the hierarchy of levels and consider entities at higher levels as real without being forced to assume new kinds of obscure metaphysical ingredients like vital forces, entelechies or souls. Higher entities do not consist of physical parts and something non-physical. All the entities being there are "physical in nature". Being physical in nature, however, does not imply that higher level entities and their properties are reducible to the sums of physical particles and their properties.

> Thus, one might want to say that higher-level entities, such as human beings, are real—as real as the entities that compose them—and at the same time reject all sorts of vitalism or dualism. (Murphy 1999, 130)

Higher level entities supervene upon their physical basis but the ontological significance of these entities is acknowledged because a comprehensive reduction to their basis is excluded.

With the distinction of reductive and non-reductive physicalism in mind it is easier to notice the conflict between (1) to (6). The conflict is to be located between the ontological primacy of the physical level on the one hand and a realist understanding of higher level entities on the other hand. Perplexities arise when we ask how higher level entities ought to be characterized ontologically. The discussion should have made clear that non-reductive physicalists have to solve a major problem of their account: The assumed dependency of higher level entities on the physical oscillates between reductionism and breaking the seal of the physical domain: Either dependency and supervenience is defined in such a way that higher level entities are identical and consequently reducible to physical ones or dependency is so weakly defined that the higher level domain gets a life on its own (Kim 1995). Then it is hard to keep the higher level domain in check within a physicalist framework. The problem for non-reductive physicalism can be stated in the form of a dilemma which is similar to the one of scientism in general: Non-reductive physicalism aims at interpreting the world of our common sense realistically. For achieving this purpose it considers mere dependency or supervenience relations on the physical as sufficient. Thereby the physical level loses in importance and definiteness. Non-reductive physicalism runs the risk to burst the physical realm altogether and thus turning into a version of liberal naturalism. If, on the contrary, the domain of the physical is kept restrictive, then non-reductive physicalism seems to slide into reductive physicalism. In what follows, we present an argument that tightens this assumption.

5. ONTOLOGICAL PHYSICALISM AND REDUCTIONISM

In various articles Kim argued at length that non-reductive physicalism is an unstable house of cards. It is a promissory note between the poles of

open mindedness and reductionist physicalism which has not been cashed out yet. Kim reminds non-reductive physicalists to be consequent:

> […] if you have already made your commitment to a version of physicalism worth the name, you must accept the reducibility of the psychological to the physical […]. (Kim 1995, 134)

Kim's argument concerns not only the mental but can be generalized for all higher level entities since the mental is just one kind of higher level entity among others (Loewer 2001b; Sparber 2005; especially Kim 2005, 52-56).

Before focusing on Kim's argument itself, something needs to be said about its premises: It is crucial for the argument that physicalists subscribe to the principle of causal closure of the physical (CCP). CCP has been stated in different constructions (e.g. Papineau 1993, 16f. and 29-32; Armstrong 1995, 38; Papineau 2001; Kim 2005, 15f.). CCP says something like the following: "At every time at which a physical state has a cause, it has a fully sufficient physical cause" (Lowe 2000, 27). Every physical state P which is caused at a certain time t has as its cause a set of other physical states existing at this certain time t, such that: (i) each of these states is a cause of P and (ii) together they are causally sufficient for P (ibid.). Physicalists "worth the name" accept CCP equally. If CCP is rejected, various kinds of non-physical entities could be accepted as causes of physical states. One major problem with this assumption is that these entities are not accessible to physics. According to most physicalists CCP is a presupposition which is part of a reasonable interpretation of our physical theories about the world. Otherwise physics could not be applied to certain domains of our reality (e.g. Armstrong 1995, 38; Beckermann 2000).

Additionally physicalists accept a supervenience relation of higher level entities to lower level ones. The supervenience relation is supposed to grant that higher level entities are 'bound' to lower level ones. The assumption of identity of higher level phenomena to physical ones is perfectly compatible with the assumption of supervenience.

If these two premises are accepted, Kim's argument briefly goes as follows (Kim 1995, Kim 2005): For the easiness of illustration, we call any arbitrary higher order state M_n and any arbitrary lower level state P_n: First, higher order states are supervenient on lower level ones. If there are two

arbitrary higher level states M_1 and M_2, then there must also be two lower level states P_1 and P_2. P_1 and P_2 are the supervenience bases of M_1 and M_2. Suppose now, that M_1 causes M_2. If this is assumed, M_1 would have to cause P_2 as well in virtue of which M_2 appears. However, because of the causal closure principle, P_2 must also be caused by P_1, while the appearance of P_2 allows M_2 to exist. As the domain of the physical is closed (according to the physicalist) M_1 cannot *by itself* cause P_2. Thus, P_1 must cause P_2. Therefore the causal chain from P_1 to P_2 and to the supervenience of M_2 is sufficient for M_2 to appear. Since a physicalist accepts the causal closure principle, the causal chain from P_1 to P_2 is not only sufficient but furthermore the only acceptable one for him. If M_1 really causes M_2, and the causal chain leading to M_2 starts with M_1's supervenience base P_1, then M_1 must be identical with P_1. Thus, taking higher level causation seriously within a physicalist ontological framework, this embraces the thesis that only reducible higher level entities possess causal efficacy. The main purpose of the argument is to show that the assumption of the supervenience relation together with CCP lead to a determinate ontological commitment. This commitment states that higher level phenomena make a difference only if—via the supervenience relation—they are identifiable and hereby reducible to physical states.

Such a conclusion follows if causal over-determination (M_1 and P_1 *each* fully *sufficiently* cause M_2 at the same time t) on the one hand and partial causation (M_1 and P_1 together *partially* cause M_2 at t) on the other hand are excluded. Assuming a (constant) causal over-determination would lead us to a highly fragmented understanding of reality because various causal histories for a single event would be equally true. Apart from this undesirable result there is no empirical evidence for it (for a further assessment of this assumption see Sparber 2005). Partial causation, on the contrary, is excluded by CCP itself, as it states that every physical event has a *sufficient* physical cause. Thus, a physicalist—by accepting CCP and some kind of supervenience relation—faces pressing ontological reasons for embracing as well the thesis that ontological relevant states at higher levels are identical and thereby reducible to the states of the bottom level of reality. CCP and supervenience seem to be intertwined insoluble with the ontological commitment that all facts are fixed on the bottom level of reality.

What CCP does not rule out are epiphenomenal states. Epiphenomenal states are higher level states which are not reducible to physical ones but which are causally powerless. As such, epiphenomenal states remain outside the realm of the physical. Nevertheless they are often considered to be no serious threat to a physicalist framework as they are causally ineffective.[7] It might, however, be asked whether epiphenomenalism truly can be dispatched so easily from physicalism. If epiphenomena are causally irrelevant, how can we know anything about them? As we have knowledge about them, they must somehow directly or indirectly affect us:

> If a thing lacks any power, if it has no possible effects, then, although it may exist, we can never have any *good reason* to believe that it exists. (Armstrong 1995, 40)

Any good reason for assuming the existence of an entity is due to its causal efficacy, by whose virtue we get knowledge of it. So either we have a good reason for assuming the existence of an entity, but then we must accept that it affects us in some way and is no epiphenomenon anymore; or it is an epiphenomenon but then we can really have no good reason for assuming its existence. Epiphenomenalism seems to be a too cheap way out for physicalism. A thoroughly coherent form of physicalism is pushed towards complete reduction.

We do not want to dwell on this issue any longer because it launches a discussion on its own. The decisive point we wanted to raise is whether it can be shown that higher level phenomena are reducible to physical entities or not. If not, we have to assume the existence of irreducible higher level phenomena. Then, a thorough physicalist ontology is unable to capture certain features of our world. Physicalism as a comprehensive ontological program has failed.[8] If it can be shown instead that no such irreducible phenomena exist, reductionism is on its move. It is our impression that many physicalists postpone the pressing answers how their

---

[7] Kim for instance takes qualia as epiphenomenal states; see Kim 2005, 22-29, and 170-173.

[8] An alternative possibility is simply to deny those entities which do not fit into a physicalist picture. Melnyk 2003, 42f., explicitly denies all entities that cannot be reduced. Whatever the merits of such a strategy are, Melnyk agrees with us that reductionism must be comprehensive and thorough.

acceptance of CCP and/or their conception of supervenience are reconcilable with an ontologically realistic understanding of a layered model of reality. If our arguments are correct they pinpoint that a non-reductive physicalist faces unavoidably conflicts between the assumptions of different levels of reality: Questions concerning the causal closure of the physical domain and the ontological overlapping or competition of higher and lower level entities call for an answer.

At this point a proponent of the reductionist program might refer to science's future success of accomplishing complete reduction. Although we are currently still remote from such a state in science's development, a glance at the history of science sheds hope: Successful programs of reduction already have been carried out and further ones will follow. Thus, it is legitimate to suppose that new forms of reduction will once be as successful as it already has been proven to be in other realms of science.[9]

We do not think that referring to successful examples in the history of science is really a convincing argument for a general reductionist outlook. There are many phenomena tenaciously resisting reduction as contemporary debates in philosophy of mind or philosophy of nature show. Nor is it satisfying to talk about 'reduction in principle' or 'reduction being possible in the long run of science'. To refer to a future point in science's progress (when the program of reduction finally will be carried out) runs a risk of becoming a mere strategy of immunisation in the face of unsuccessful attempts of reduction. The same holds for the claim that reduction is possible in principle, but cannot be executed in practice because the entities which are to be reduced are too complex. The physicist Falkenburg plausibly demands that successful reductionism has to be carried out in a double way: It needs to analyse higher level entities into physical ones (top down reduction) on the one hand and it has to be shown how higher level phenomena result from its physical constituents (bottom up reduction) on the other hand (Falkenburg 2006, 61-68).

---

[9]   See, for instance, Kim 2005, 68, citing McLaughlin 1992.

6. CONCLUSION

Naturalism is a very popular philosophical position. We share its underlying conviction that modern sciences help us to see many things clearer, also in the field of ontology. But the problem is that there is no direct way from the subject matter of sciences to ontology. The aim of our paper was not to present new arguments in favour of naturalism or against it. Its aim was to map the landscape of the current discussion of ontological naturalism and to portray the ways a naturalist might want to go within it.

We pointed out that naturalism has no clear concept of science. Therefore it is hard to tell from which sciences naturalists should or could derive their ontology. Even if a list of acceptable sciences were provided, serious questions would remain: How do the ontological implications of the accepted sciences relate to each other? And do some sciences have a higher priority in determining ontology than others?

According to our analysis the most promising way for naturalists to elaborate a science-based, uniform, and coherent ontology is physicalism. Many of the contemporary physicalistic positions, however, accept CCP and supervenience—the core assumptions of physicalism—on the one hand but reject reductionism on the other hand. We do not see how this wish-list of non-reductive physicalism can be brought together consistently. A physicalist is committed to reductionism, as Loewer[10] unwittingly summarizes our conclusion:

> […] philosophers true to their physicalism will have to swallow reductionism. Those who find reductionism impossible to swallow will have to find a way of living without physicalism. (Loewer 2001b, 315)

Some naturalists might have the impression to face another 'straw man argument' of naturalism being fought here. But we aimed at interpreting naturalism benevolently. We neither rejected its allegiance to science nor did we claim naturalism to be a failure as a philosophical program. We rather wanted to press the naturalist to consider more carefully the

---

[10] Loewer himself wants to avoid this conclusion—against his line of argument see Sparber 2005.

ontological commitments she enters by taking science as authoritative guidance for philosophy in general and ontology in particular.

We do not know where science will lead us and whether naturalism will have a better standing in the future. Contemporary naturalism, however, carries a heavy burden: If it wants to be successful, it either has to show how reductionism is possible, or it has to point out convincing ways for manoeuvring between reductionism and too liberal versions of naturalism.

7. REFERENCES

Armstrong, D. 1995: "Naturalism, Materialism, and First Philosophy." In: Moser, P. K. and Trout J. D. (eds.) *Contemporary Materialism*. New York: Routledge, 35-47.

Beckermann, A. 2000: "Ein Argument für den Physikalismus." In. Keil, G. und Schnädelbach, H. (eds.) *Naturalismus*. Frankfurt am Main: Suhrkamp, 128-143.

Clayton, P. 2004: *Mind and Emergence. From Quantum to Consciousness*. Oxford: Oxford University Press.

Crane, T. and Mellor, H. 1995: "There is no question of physicalism." In: Moser, P. K. and Trout J. D. (eds.) *Contemporary Materialism*. New York: Routledge, 65-89.

Dupré, J. 2004: "The Miracle of Monism." In: De Caro, M. and Macarthur, D. (eds.) *Naturalism in Question*. Cambridge, MA: Harvard University Press, 36-58.

Falkenburg B. 2006: "Was heißt es, determiniert zu sein? Grenzen naturwissenschaftlicher Erklärung." In: Sturma, D. (ed.) *Philosophie und Neurowissenschaften*. Frankfurt am Main: Suhrkamp, 43-74.

Forrest, P. 1996: *God without the supernatural: A defense of scientific theism*. Ithaca: Cornell University Press.

Hüttemann, A. and Papineau, D. 2005: "Physicalism decomposed." *Analysis 65*, 33-39.

Keil, G. 2000: "Naturalism in the philosophy of mind—and what is wrong with it." In: Nannini, S. and Sandkühler, H. J. (eds.) *Naturalism in the Cognitive Sciences and the Philosophy of Mind*. Frankfurt am Main: Peter Lang, 145-155.

Keil, G. and Schnädelbach, H. 2000: "Naturalismus." In: Keil, G. and Schnädelbach, H. (eds.) *Naturalismus. Philosophische Beiträge*. Frankfurt am Main:Suhrkamp, 7-45.

Kim, J. 1995: "The Myth of Nonreductive Materialism." In: Moser, P. K. and Trout J. D. (eds.) *Contemporary Materialism*. New York: Routledge, 133-149.

Kim, J. 2005: *Physicalism or something near enough*. Princeton, NJ: Princeton University Press.

Koppelberg, D. 2000: "Forms of Naturalism in the Philosophy of Mind." In: Nannini, S. and Sandkühler, H. J. (eds.) *Naturalism in the Cognitive Sciences and the Philosophy of Mind*. Frankfurt am Main: Suhrkamp, 11-29.

Kornblith, H. 1994: "Naturalism: Both Metaphysical and Epistemological." *Midwest Studies in Philosophy* 29, 39-52.

Loewer, B. 2001a: "From Physics to Physicalism." In: Loewer, B. and Gillet, C. (eds.) *Physicalism and its Discontents*. Cambridge: Cambridge University Press, 37-56.

Loewer, B. 2001b: "Review of Jaegwon Kim, Mind in a Physical World. An essay on the mind-body problem and mental causation." *Journal of Philosophy* 98, 315-324.

Löffler, W. 1999: "Naturalisierungsprogramme und ihre methodologischen Grenzen." In: Quitterer, J. and Runggaldier, E. (eds.) *Der neue Naturalismus*. Köln: Kohlhammer, 30-76.

Lowe, J. E. 2000: *Introduction to the philosophy of mind*. Cambridge: Cambridge University Press.

McLaughlin, B. 1992: "The Rise and Fall of British Emergentism." In: Beckermann, A.; Flohr, H. and Kim, J. (eds*.) Emergence or reduction?*: *Essays on the prospects of Nonreductive Physicalism*. Berlin, New York: De Gruyter, 49-93.

Melnyk, A. 2003: *A physicalist manifesto: thoroughly modern materialism*. Cambridge: Cambridge University Press.

Murphy, N. 1999: "Nonreductive Physicalism. Philosophical Issues." In: Brown, W. S., Murphy, N. and Malony, N. (eds.) *Whatever Happened to the Soul? Scientific and Theological Portraits of Human Nature*. Minneapolis, MN: Fortress Press, 127-148.

Papineau, D. 1993: *Philosophical Naturalism*. Oxford: Oxford University Press.

Papineau, D. 2001: "The Rise of Physicalism." In: Loewer, B. and Gillet, C. (eds.) *Physicalism and its discontents*. Cambridge: Cambridge University Press, 3-36.

Pettit, P. 1993: "A definition of Physicalism." *Analysis* 53, 213-223.

Poland, J. 1994: *Physicalism. The Philosophical Foundations*. Oxford: Clarendon Press.

Rea, M. 2002: *World Without Design: The Ontological Consequences of Naturalism.* Oxford: Oxford University Press.

Sparber, G. 2005: "Counterfactual overdetermination vs. the causal exclusion problem." *History and Philosophy of the Life Sciences* 27, 479-490.

Stroud, B. 2004: "The Charm of Naturalism." In: De Caro, M. and Macarthur, D. (eds.) *Naturalism in Question*. Cambridge, MA: Harvard University Press, 21-35.

Sukopp T. 2006: *Naturalismus. Kritik und Verteidigung erkenntnistheoretischer Positionen.* Frankfurt am Main: Ontos.

Tetens, H. 2000: "Der gemäßigte Naturalismus der Wissenschaften." In: Keil, G. and Schnädelbach, H. (eds.) *Naturalismus. Philosophische Beiträge*. Frankfurt am Main: Suhrkamp, 273-288.

Reason, Red in Tooth and Claw: Naturalising Enlightenment Thinking

Konrad Talmont-Kaminski
Marie Curie-Sklodowska University

1. INTRODUCTION

The Enlightenment raised reason to be the main human virtue. However, the picture it had of reason was not borne out by time and investigation: the twentieth century has seen a growing cynicism concerning Enlightenment values, leading to various forms of anti-intellectualism, be it nihilist or fundamentalist. To a great degree, much the same process can be seen to have taken place in philosophy:

> After Popper, through Kuhn, Feyerabend, and all the others, the appeal to decisions by the scientific community widens rapidly—so rapidly that all of these latter have been accused of abandoning reason. Why? Only because of the tacit assumption that what cannot be reduced to logical method is nonrational. But this consequence is instead better taken as a reductio of this conception of rationality. (Hooker 1995, 27-28)

It is in this context that naturalism can make its most valuable contribution. It does so by accepting that the traditional view of rationality is bankrupt but, instead of abandoning Enlightenment ideals, shows how they can be salvaged around the basis of a naturalised rationality. This naturalised rationality is understood in the context of our biological, evolutionary inheritance and, lacking the hubris of the traditional view, offers not the guarantee of perfection but the possibility of progress. In doing so, however, it avoids the objections which brought down the traditional view while still remaining robust enough to provide reason to think that we can transcend our existing limitations. As such, it offers a viable humanist response to the anti-intellectualisms of nihilism and fundamentalism.

2. NATURALISM WITHIN THE ENLIGHTENMENT TRADITION

The twentieth century has been witness to an on-going onslaught upon the Enlightenment tradition. This has been the case both upon the grand stage of world events—rocked as they were time and again by factions espousing value systems antithetical to Enlightenment thinking—and the smaller stage of intellectual history, upon which the Enlightenment virtue of reason has been under attack from a number of directions. Naturalism, in the form it takes upon this smaller stage, can be usefully understood as the attempt to salvage what can be saved from the traditional Enlightenment values following the critiques they faced throughout the twentieth century.

Pascal, famously, spoke of humans as thinking reeds (Pascal 1669, pensée 347)—weak in flesh but magnificent in mind. For Pascal, just like for the other Enlightenment thinkers such as Voltaire, Diderot, Rousseau and Montesquieu, reason was the one human trait that allowed us to transcend our limitations, the one human trait that made progress a salve for the human condition. In their view of human reason, the Enlightenment thinkers showed more than just a slight tendency towards dualism—with reason belonging to some sphere largely independent of the reed-like human bodies. This view was understandable both historically and conceptually. Understandable historically; because Pascal and the others, despite the many differences between them and Descartes, to a great degree followed in Descartes' dualist footsteps and, partly through him, were steeped in the profoundly dualist Christian intellectual tradition. Understandable conceptually; since it was hard to see how reason could provide a way of transcending human frailties if it were intimately engaged with them.

By the beginning of the twentieth century, Enlightenment reason came to be understood as embodied in science and logic—these two being seen as the engines of human progress and as responsible for the great changes that took place in human well-being and capabilities in the previous century. Infamously, those engines carried us to the fields of Passchendaele, the ovens of Treblinka and the skies above Hiroshima—leading to a growing cynicism regarding the virtue of reason and science (Horkheimer and Adorno 1947). On the intellectual scene, the early-

twentieth-century inheritors of the Enlightenment French were, foremost among others, the Viennese logical positivists. They, with their generally socially progressive ideals, were among the first to suffer due to the darkening character of European politics in the inter-war period and were forced to flee, primarily to America. There, their views were highly influential but also came under gradually more corrosive versions of criticism that ranged from the logical positivists' own awareness of the limitations of their project to espousals of a total rejection of their ideal of reason expressed by Rorty and many others that could be broadly characterised as postmodernists. Some of positivism's most extreme critics see positivism and fascist ideologies as part and parcel, the actual historical relation between those views being something of an embarrassing difficulty for that 'discourse'—as is the actual intellectual genealogy of recent anti-Enlightenment philosophies, reaching back as it does to the pro-fascist literati of the 1930s (Wolin 2004).

Of course, neither the Viennese nor the other philosophers who saw themselves as carrying on the Enlightenment tradition were of one voice on any of a number of issues. Still, both they and their critics saw the strength or weakness of reason as turning on its intimate relationship with logic; whose sovereignty was the only guarantee that reason could be free of the baser aspects of our nature.

Brown (1988) characterises this logic-based view of rationality as having three features:

(1) Rational solutions should be universal.
(2) Rationally acceptable conclusions must follow with necessity from the information given.
(3) The rationality of the conclusions is determined by whether they conform to the appropriate rules. (after Hooker 1991, 44f.)

The view is an understandably attractive one. On the one hand, all three features may be thought to be realistic so long as the rules of rational reasoning could be simply read off logical relations. On the other hand, the form of reason characterised by them appears to offer us the opportunity to escape the contingencies of human foibles. Together, the view makes

concrete the way to achieve the transcendence the Enlightenment thinkers dreamed of.

One of the attempts to arrive at a model of such rationality took the form of the search for an inductive logic (Keynes 1921, Carnap 1950). For all the efforts that have been put into the search, however, Hume's problem remains as intransigent as ever. Yet, at the same time, the ampliative inferences we have continued to make—without a basis in any satisfactory theoretical understanding of induction—have continued to succeed at a rate that, in light of the problem of induction, must look miraculous.

Hooker (1991) recognises that much of the criticism raised against the traditional conception of reason, which depended to a great degree upon the sovereignty of logic, was apposite. The long litany of critiques that have dogged logic-based accounts of rationality over the major part of the twentieth century (Popper 1935, Goodman 1955, Quine 1960, Kuhn 1962, Feyerabend 1975 and many others) together with the failure to answer Hume's original problem has done grievous damage to the notion of a universal, logically necessary, rule-governed rationality. While some philosophers try to forge such a logicist account regardless, many feel compelled by the failure to reject Enlightenment ideals and to accept some sort of epistemic nihilism. Hooker, as well as many other naturalists, would seek something like a middle path between those two options:

> I accept that Feyerabend is essentially correct in his insistence that the Western project for reason, as it is encapsulated along the rationalism-empiricism axis, is in serious difficulty and requires some superseding conception or other. But while Feyerabend's response has been to question, in an increasingly radical way, the transcendence project which it theorizes I shall argue that there is an alternative: accept the transcendence project and re-theorize the nature of reason and the life of reason. (Hooker 1991, 44)

This, then, is the naturalist position I wish to espouse—to continue with the Enlightenment project while rejecting the traditional view of rationality. This rejection goes much deeper than the particularities of the logicist view of reason that was constructed in the early twentieth century, however: the claim being made is that the Enlightenment's weakness was that it did not go far enough in its humanism by failing to bring reason down to the human level. In retaining the Cartesian view of reason as something semi-

divine, Enlightenment thinking was guilty of hubris (not in the face of the gods but of nature, and of human nature in particular). With the gift of hindsight, the fall was inevitable. Given such a diagnosis, naturalism must do more than just replace the logicist view of reason with another account that claims for it unpossessed virtues. The foremost among them have to be claims of certainty that, as the Enlightenment's critics have been right to point out, have provided a rationalisation for the stifling of discussion after a 'rational' conclusion was reached. Of course, the sort of universal fallibilism that results from recognising our incapacity to achieve certainty is something that goes back to Peirce's (1868a, 1868b) critique of Descartes and has, thankfully, come to be generally accepted today both within and outside of naturalist circles. Still, for their fallibilism to be more than scepticism and to be able to continue with the transcendence project, the naturalists need to say something more positive about reason.

3. A PLETHORA OF NATURALISMS

Since Dewey's (1925) espousal of naturalism, a richly branching family of various positions that go by that name has evolved. Indeed, when one considers all these positions it is hard to identify anything that they have in common beyond a generally favourable attitude to science. At times, in fact, nothing more seems requisite for a view to be deemed naturalistic. Just how wide the scope is can perhaps be gauged when it is considered that Frank Jackson's (1997) defence of conceptual analysis is often considered a good example of naturalism, even while Millikan (2005) and Bishop and Trout (2005) see a robustly critical attitude to conceptual analysis as essential to naturalism. A term is only as useful as the class of things it identifies and, in the case of 'naturalism', the breadth of the common meaning is such as to have made it difficult to have a focussed discussion as anything like a core of the position has been obscured by vagueness. A case study of the difficulties this has caused is provided by the critical articles collected by De Caro and Macarthur (2004), which all too often end up being aimed at nothing more than some vague scientism.

Characterising naturalism as the rejection of logic-based accounts of reason together with a retention of Enlightenment ideals still leaves naturalism a broad church, but it helps to focus the debate around a clear

position. Doing so, however, entails taking sides in naturalism's internecine struggles, particularly that over the question whether naturalism is to be primarily characterised in terms of its methodological or metaphysical claims. The distinction may be drawn (see, for example, Audi 1995, 517-518 or Knowles 2003) as follows:

Metaphysical naturalism takes as its starting point the assertion that only certain (naturalist) metaphysical claims are acceptable and that other (supernatural) claims are to be rejected. Most often the underlying metaphysical distinction is made on the basis of the ontology used by science by claiming that only entities recognised by science should be called upon by naturalist philosophy.

Methodological naturalism, on the other hand, has as its starting point the assertion that only certain kinds of epistemic methods can be rational. Again, it is typically—though not necessarily—science that provides the measure for what methods are acceptable. The relevant alternative is provided by the a priori/a posteriori distinction, with aposteriorist methodological naturalism claiming that most or all rational epistemic methods are only justifiable a posteriori (for example Haack 1993, 118ff).

To avoid a possible confusion it is important to note that the distinction between a metaphysical and a methodological naturalism appears also in a different context (for example in Plantinga 1996), in which that particular "methodological naturalism" is merely a weaker version of the metaphysical position in so far as the scientific ontology is accepted by it as just a working hypothesis. The specific context in which that distinction was originally made (possibly by De Vries 1986) is that of the intellectually bankrupt propagandising of the supposed virtues of various pseudo-scientific transmogrifications of creationism. The overall confused state of creationism's defenders can be seen in how that distinction is doubly defective: both because it distinguishes not between kinds of naturalist views but how they are held and because—given the general fallibilism of both science and naturalism—naturalist views can only ever be held as working hypotheses, rendering "metaphysical naturalism" an obviously self-contradictory straw man and leaving the so-called "methodological naturalism" the only live alternative. I will say no more of that distinction.

The naturalist position I see Hooker taking (1997, 52), and the one I would argue for, is strong aposteriorist methodological naturalism, i.e. the view that *all* epistemic methods can at best be provided with a posteriori justifications only. In effect, first philosophy is abandoned and epistemic methods are recognised to be context-dependent, theory-laden and, of course, fallible both in their justification and their application. By prioritising the methodology over the metaphysics, this approach to naturalism to a large degree forestalls the questions of reductionism that plague metaphysical naturalism. By claiming that epistemic methods can only be justified a posteriori, this form of naturalism picks a middle path between the traditional position and a scientistic naturalism that would claim that science is the sole source of rational epistemic methods. The rejection of the traditional position that at least some epistemic methods are justified a priori seems to be necessitated by the failure of more than two-hundred-years-worth of attempts to find a solution to Hume's problem. Further support for this view can be seen to come from Quine's (1951) refutation of empiricist dogmas, the relevant implications having been pursued by Putnam (1976). On the other hand, the scientistic position seems to be not just insufficiently motivated but, more importantly, to run counter to what science tells us about the continuum of epistemic methods (Campbell 1974) that runs from the simple chemotaxis of single-celled organisms such as the paramecium, through such everyday uses of perception as looking both ways before you cross the road, to the ever-growing family of highly specific methods used (and tested) in science. Of course, Hooker and I are hardly the only supporters of the naturalist position under review (Haack 1993 and Rescher 1977 provide just two possible additional examples) and much work has already been done to motivate this methodological version of naturalism.

The move to give up on first philosophy brings both riches and troubles. The troubles come in the form of the knowledge provided by the science and whatever other effective epistemic methods humanity uses, which, given the rejection of a method foundationalism—or *foundationalism* as Haack (1993, 186) refers to it—naturalists are free to engage in coming to understand what methods are rational. The troubles are primarily those which affect any non-foundationalist position; particularly, the issue of justification that has already been raised. The naturalist must hope that

their riches are sufficient to buy their way out of trouble. I will consider both in turn.

4. THE RICHNESS OF NATURALISM

Naturalism is, by its very nature, interdisciplinary. It does not have any a priori qualms about the relevance of how we reason to the question of how we should reason. As such, it can draw upon the lessons provided by the history of science as well as the sociology of science, much in the way that Kuhn (1962) suggested it. A historicised and sociologised philosophy of science has been a major step forward, however naturalism goes further. At the same time as it accepts the data from studies of scientific practice, it also learns from how living beings in general manage to deal with their environment, requiring it to draw upon the work of biologists and neuroscientists as well as that of other scientists. In effect, the sciences provide not only a useful case study for naturalised philosophy but are also valuable partners in the enterprise of seeking to understand how knowledge is actually obtained and used in the world, be it by humans or by their evolutionary predecessors. The approach that has been pursued for decades now by a number of evolutionary epistemologists (Campbell 1974, Plotkin 1982, Callebaut and Pinxten 1987) is just one example of this kind of interdisciplinarity. All of these disciplines provide a great wealth of knowledge about what reasoning and its evolutionary precursors are actually like.

   Naturalised approaches to reasoning are also interdisciplinary in another respect. By considering reason in the concrete context of actual reasoning beings rather than as an abstract set of rules and relations, naturalism forces together epistemology and philosophy of mind, making it necessary to consider in close connection to each other such things as epistemic methods, rational decisions and other aspects of philosophy of mind. The result of the bringing together of what had seemed like disparate philosophical and scientific disciplines has been the embodied, situated cognitive science that has come of age during the last decade of the previous century and which is being developed by Andy Clark (1996), Dan Dennett (1997), David Chalmers (1996), Mark Bickhard and Loren Terveen (1995) and Shaun Gallagher (2005) among many others. Clark's

work provides a good example of this kind of thoroughly naturalised and unashamedly multidisciplinary approach. Thus, his analysis of the role played by perception moves beyond philosophical standards that assumed the construction of a complete and neutral model of the environment, and instead learns from neurological and other empirical studies that perception is highly constructive and closely tied to action, giving us access, at the right time, to information which is adequate and in the appropriate form to make the decisions that need to be made at that very point. Such a focus on practical commitment is one that can be seen to run through naturalism.

At this point it might seem that the tag of scientism is one that well fits naturalism, given that it is various scientific disciplines that have been repeatedly called upon. To see why this is not a fair accusation it is necessary to consider further the relationship between science and naturalised philosophy of the type I argue for. The heart of the matter is the fact that the definition of this naturalism is to be found not in the acceptance of scientific methods, however they were to be demarcated, but in the rejection of the possibility of having any a priori justification for epistemic methods. Given the rejection of the a priori, in so far as any methods are to be favoured, such status can only be given on the basis of the observed results of applying those methods, be they "scientific" or otherwise. This means that, in so far as naturalism views science favourably, it does so only a posteriori—having seen that scientific methods do seem to provide an effective way to come to know the world we live in. Exactly the same is true for non-scientific methods, be they everyday practices people apply when crossing the street or choosing fruit at a grocer's, or be the activity under consideration literature or poetry, both of which I would argue have a cognitive, epistemic role. Given that naturalism does not see science as a set of universal methods, scientific activity can be seen as continuous with other human activities, different only in the degree of institutionalisation, formalisation, self-criticism, and self-awareness regarding the particular context-dependent methods it uses.

Considering the richly intertwined continuum of evolutionary developments in the means used to observe the environment that science reveals, any scientism that seeks to sharply distinguish between scientific methods and other means living beings use to make their way in the world is not just false but unscientific. Thus, the continuum of human methods

must be seen as part of the continuum of the methods used by living beings in general. As such, human science, human knowledge and—perhaps most importantly—human reason must likewise be seen as part of that continuum. The precise path this continuum follows can be traced both in terms of the kinds of distinctions organisms are capable of making—from the identification of the slope of a sugar gradient made by a paramecium, through the human ability to identify someone they know merely by hearing their footsteps, to the detection of a subatomic particle in a particle accelerator—as well as in terms of the kinds of models organisms use to organise their knowledge—from the models pragmatically implicit in the reactions of that paramecium, through the explicit understanding people have of their surrounds, to the purely mathematical models used when dealing with the counterintuitive nature of the quantum world. One detailed discussion of just how these various abilities build upon each other is provided by Bickhard (2003) in his examination of how this process is tied to the emergence of reference, which is normally seen as a distinguishing mark of real cognition.

5. THE TROUBLES NATURALISM FACES

Naturalism's openness to various methods also has its downside. By accepting the input of other disciplines and epistemic methods, naturalist accounts of reason are forced to somehow deal with the fact that these various sources often proscribe differing courses of action. This is as true when comparing clinical medicine with homeopathy, or people's fast and frugal heuristics (Gigerenzer 2000) with Bayesian rules of inference, as when the comparison is between how a chemist and a paramecium would locate concentrations of dissolved sugars. An uncritical pluralism regarding methods is quite misplaced here. To use Peirce's (1877) phrase, "real and living doubt" will have real consequences in terms of the sorts of ways one will act and, although in discussion we may eschew commitment, we cannot avoid it in action. At the same time uncritical conventionalism, though quite capable of guiding actions, will not be capable of improving upon the status quo. In effect, the naturalist must find a way between conventions and anarchy.

Invoking Peirce in this context is most appropriate, as a number of prominent naturalists who have grappled with these issues (Haack 1993, Hooker 1991, Hacking 1983, Hookway 2000, Rescher 2005) have found his pragmatism to be capable of providing the right framework for working towards an answer. Already we have seen something of Peirce's view in the synechist refusal to draw sharp distinctions between science and other epistemic methods. In this instance, however, the vital aspect of Peirce's thought is his (1878) insistence that "the whole function of thought is to produce habits of action". Hooker presents the point this way:

> [A] shift to an alternative paradigm of acceptance as an act of practical commitment leads to a decision theoretic epistemology where consequences play a key role, acceptance strategies can be subject to selective development, and there is no requirement for foundations. (Hooker 1991, 63)

The approach offers the hope that it is possible to get away from a foundational model, in this case supposedly built up on a bedrock of a priori methods, without giving up on the transcendence project, as it is now the practical commitment which is to underwrite the necessary normativity. The question of normativity is vital for naturalism if the so-called naturalistic fallacy (Moore 1903) is to be shown to be no fallacy at all. That is, indeed, the aim of the neo-Peircean, biologically inspired project to develop a naturalist account of function that is pursued by Hooker, Bickhard (2003) and John Collier (2000) among others (and which is actually quite different from that pursued by Millikan (1984); who is also influenced by Peirce and biology, of course). In fact, given failure to provide a robust basis for normativity in anything other than the brute facts of our embodied selves being situated in our environment may, I would argue, lead to the conclusion that the actual fallacy was an anti-naturalist one. At the same time, it is important to note naturalist arguments to the effect that normativity is not actually needed, a recent influential argument to that effect having been put forward by Jonathan Knowles (2003).

Basing our understanding of methodology on a posteriori methods allows us to dull the edge of the objections that dealt the killing stroke to the traditional view. Treating the epistemic methods we use as open to development in much the same way as our beliefs about the world we live in (i.e. thinking that methods are essentially beliefs about how we go on

about finding out about our world) provides a way to internalise Kuhn's and other's objections to the logic-based account of rationality. This, immediately, gives a strategy to dealing with Hume—that strategy being to deal how to live with his objection rather than trying to show it to have been somehow mistaken. Counter to Hume's seeming assumption, rationality turns out to be situated not in some universal method but in something like the humble habits that he, himself, fell back upon. There can be no guarantees that our methods will not fail us the next time we apply them. Yet, given the brute fact of our own (fragile) existence and the unavoidability of action, we are forced to chose on the basis of what understanding we hope we do possess. The sceptic has nothing to offer us in our need since, if we accepted their advice, we should have no basis to make any of the decisions that are forced upon us. Thankfully, Hume does not show that our methods will always fail in all possible worlds, but, rather, shows than there are no methods that will work across all of them. In its modesty, naturalism acknowledges this inevitability and only claims to seek methods that will be adequate in certain limited contexts. This response seems, in fact, to be very much in line with the kind of naturalism that Kemp Smith (1905a, 1905b) originally and H.O. Mounce (1999) recently see as lying behind Hume's arguments and his own not-so-sceptical responses.

6. NATURALIST'S TRANSCENDENCE

At this point we are in the position to try and draw up a characterisation of the kind of naturalised rationality that the account I pursue leads to. Thinking back to Brown's (1988) characterisation of logic-based rationality we can juxtapose to it the features a naturalised rationality has:

(1) Rational solutions (or, more appropriately, methods) need not be universal, but only have to be effective in the appropriate limited contexts.

(2) Rationally acceptable conclusions (of psychological inferential processes) do not have to follow necessarily from the information given, as acceptance is not to be understood

in terms of a formal relationship between propositions but in terms of a practical commitment.

(3) The rationality of the conclusions (of psychological inferential processes) is not necessarily determined by whether they conform to the appropriate rules, indeed the primary focus is removed from the conclusions and placed upon the actions that are taken on the basis of the practical acceptance of beliefs and methods that are all subject to further criticism and development.

This vision of rationality fits into the evolutionary synthesis (Huxley 1942) that first united various elements of biology and which is now bringing together an ever broader range of sciences. As such, it is a world away from the traditional Enlightenment view of rationality as a spark of something close to divinity trapped within brute beasts. Given that, and given the kind of limitations the new view of rationality places upon it, what prospect for transcendence can remain?

One of the core ideas within the evolutionary synthesis is that of emergence: the emergence of life out of inanimate matter, the emergence of complex life-forms from single celled organisms and, in our case, the emergence of rationality out of the simple forms of goal-seeking behaviour. Without going into the controversial details of emergence it still ought to be clear how significant that concept is to naturalism's metaphysical aspects, providing as it does scope for explaining the existence of mental entities without calling for either eliminative reductionism or dualism. Emergence is significant in another way—it shows how transcendence may lead to something which is totally new. This obviates the need for reference to some pre-existing standard or goal, marking the underlying notion of progress as one that is moving us away from ignorance and parochialism rather than towards omniscience and universality. Thus, the kind of transcendence that a naturalist can see rationality providing is one that stepwise moves us beyond our existing limitations and thereby reveals further limitations to be transcended. As such, reason becomes not just the engine of progress but also one of its main products, and science is not thought of as contrary to our biological inheritance but as emergent from it. Transcendence, on this picture, is

achieved through the sometimes inventive and sometimes methodical application of our limited abilities, context-dependent methods and imperfect knowledge to augment our abilities, extend the reach of our methods and to add to our knowledge, all without ever negating the underlying boundedness of reason. This means that being rational entails neither being safe from human frailties nor not in need to further improvement. Still, the value of the kind of transcendence that a naturalist can offer lies simply in allowing us to do more things than we were able to previously and, at the same time, to understand our situation better than we were before, all the while making clear that we can never rest on our laurels. And that this is a worthwhile viewpoint can be seen all the more starkly against the background of the nihilisms of the last century as well as of the fanaticisms of the current century's opening years.

7. REFERENCES

Audi, R. (ed.) 1995: *The Cambridge Dictionary of Philosophy*. Cambridge: Cambridge University Press.

Bickhard, M. and Loren T. 1995: *Foundational Issues in Artificial Intelligence and Cognitive Science: Impasse and Solution*. New York: Elsevier Scientific.

Bickhard, M. 2003: "Process and Emergence: Normative Function and Representation." In: J. Seibt (ed.) *Process Theories: Crossdisciplinary Studies in Dynamic Categories*. Dordrecht: Kluwer Academic, 121-156.

Bishop, M. and Trout, J. D. 2005: *Epistemology and the Psychology of Human Judgment*. New York: Oxford University Press.

Callebaut, W. and Rik P. (eds.) 1987: *Evolutionary Epistemology: A Multiparadigm Program With a Complete Evolutionary Epistemology Bibliography*. Dordrecht: Reidel (Synthese Library 190).

Campbell, D. T. 1974: "Evolutionary Epistemology." In: P. A. Schilpp (ed.) *The philosophy of Karl R. Popper*. LaSalle, IL: Open Court, 412-463.

Carnap, R. 1950: *Logical Foundations of Probability*. Chicago: University of Chicago Press.

Chalmers, D. 1996: *The Conscious Mind*. Oxford: Oxford University Press.

Clark, A. 1996: *Being There*. Cambridge, MA: MIT Press.

Collier, J. 2000: "Autonomy and Process Closure as the Basis for Functionality." In: Chandler, J. L. R. and van de Vijver, G. (eds.) *Closure: Emergent Organizations and their Dynamics*, 280-291.

De Caro, M. and Macarthur, D. (eds.) 2004: *Naturalism in Question*. Cambridge, MA: Harvard University Press.

Dennett, D. 1997. *Kinds of Minds: Towards an Understanding of Consciousness*. New York: Basic Books.

Dewey, J. 1925: *Experience and Nature*. Chicago: Open Court.

Feyerabend, P. 1975: *Against Method*. London: New Left Books.

Gallagher, S. 2005: *How the Body Shapes the Mind*. Oxford: Oxford University Press.

Gigerenzer, G. 2000: *Adaptive Thinking*. New York: Oxford University Press.

Goodson, N. 1955: *Fact, Fiction, and Forecast*. Cambridge, MA: Harvard University Press.

Haack, S. 1993: *Evidence and Inquiry*. Oxford: Blackwell.

Hacking, I. 1983: *Representing and Intervening*. Cambridge: Cambridge University Press.

Hooker, C. 1991: "Between formalism and anarchism." In: Munévar, G. (ed.) *Beyond Reason*. Dordrecht: Kluver Academic.

Hooker, C. 1995: *Reason, Regulation, and Realism*. Albany, NY: State University of New York Press.

Horkheimer, M. and Adorno, T. 1947: *Dialektik der Aufklärung*. Amsterdam: Querido Verlag.

Hookway, C. 2000: *Truth, Rationality and Pragmatism*. Oxford: Oxford University Press.

Hume, D. 1739/2000: *A Treatise of Human Nature*. Edited by Norton, D. F. and Norton, M. J.. Oxford: Oxford University Press.

Huxley, J. 1942: *Evolution: The Modern Synthesis*. London: George Allen and Unwin.

Jackson, F. 1997: *From Metaphysics to Ethics*. Oxford: Claredon Press.

Kemp Smith, N. 1905a: "The Naturalism of Hume I." *Mind* 14, 149-173.

Kemp Smith, N. 1905b: "The Naturalism of Hume II." *Mind* 14, 335-347.

Keynes, J. M. 1921: *Treatise on Probability*. London: Macmillan & Co.

Knowles, J. 2003: *Norms, Naturalism and Epistemology: The Case for Science without Norms*. London: Palgrave Macmillan.

Kuhn, T. 1962: *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Millikan, R. G. 1984: *Language, Thought, and Other Biological Categories*. Cambridge, MA: MIT Press.

Millikan, R. G. 2005: *Language: A Biological Model*. Oxford: Oxford University Press.

Moore, G. E. 1903: *Principia Ethica*. Cambridge: Cambridge University Press.

Mounce, H. O. 1999: *Hume's Naturalism*. London: Routledge.

Pascal, B. 1669/1962: *Pensées*. Edited by Lafuma, L. Paris: Les Libraires Associés.

Peirce, C. S. 1868a: "Questions Concerning Certain Faculties Claimed for Man." *Journal of Speculative Philosophy* 2, 103-114.

Peirce, C. S. 1868b: "Some Consequences of Four Incapacities." *Journal of Speculative Philosophy* 2, 140-157.

Peirce, C. S. 1877: "The Fixation of Belief." *Popular Science Monthly* 12, November 1877, 1-15.

Peirce, C. S. 1878: "How to Make Our Ideas Clear." *Popular Science Monthly* 12, January 1878, 286-302.

Plantinga, A. 1996: "Methodological naturalism?" In van der Meer, J. (ed.) *Facets of Faith and Science*. Lanham, MA: University Press of America, 177-221.

Plotkin, H. C. (ed.) 1982: *Learning, Development, and Culture: Essays in Evolutionary Epistemology*. New York: John Wiley & Sons.

Popper, K. 1935: *Logik der Forschung*. Vienna: Julius Springer.

Putnam, H 1976: "'Two Dogmas' Revisited." In: G. Ryle (ed.) *Contemporary Aspects of Philosophy*. London: Oriel Press, 202-213.

Quine, W. V. O. 1951: "Two dogmas of empiricism." *Philosophical Review* 60, 20-43.

Quine, W. V. O. 1960: *Word and Object*. Cambridge, MA: MIT Press.

Rescher, N. 1977: *Methodological Pragmatism*. New York: New York University Press.

Rescher, N. 2005: *Reason and Reality*. Lanham, MD: Rowman & Littlefield.

Vries, P. de 1986: "Naturalism in the natural sciences." *Christian Scholar's Review* 15, 388-396.

Wolin, R. 2004: *The Seduction of Unreason: The Intellectual Romance with Fascism from Nietzsche to Postmodernism*. Princeton, NJ: Princeton University Press.

Part III

The Human Person

Naturalism and the First-Person Perspective

Lynne Rudder Baker
University of Massachusetts Amherst

The first-person perspective is a challenge to naturalism. Naturalistic theories are relentlessly third-personal. The first-person perspective is, well, first-personal; it is the perspective from which one thinks of oneself *as* oneself* without the aid of any third-person name, description, demonstrative or other referential device. The exercise of the capacity to think of oneself in this first-personal way is the necessary condition of all our self-knowledge, indeed of all our self-consciousness. As important as the first-person perspective is, many philosophers have not appreciated the force of the data from the first-person perspective, and suppose that the first-person perspective presents no particular problems for the naturalizing philosopher. For example, Ned Block commented, "It is of course [phenomenal] consciousness rather than...self-consciousness that has seemed such a scientific mystery" (Block 1995, 230). And David Chalmers says that self-consciousness is one of those psychological states that "pose no deep metaphysical enigmas" (Chalmers 1996, 24).

In contrast to those with such casual attitudes toward self-consciousness stands Thomas Metzinger, a naturalistic philosopher who sees the complexity of self-consciousness and treats it in detail. Some years ago, I wrote an article, "The First-Person Perspective: A Test for Naturalism" (Baker 1998), in which I presented the first-person perspective as a challenge to naturalism—at least for the robust form of reductive naturalism that aims to provide accounts of all phenomena in terms accepted by the natural sciences. Metzinger has taken up this challenge, both in his article "Phenomenal Transparency and Cognitive Self-Reference" (Metzinger 2003a) and in his book *Being No One: The Self-Model Theory of Subjectivity* (Metzinger 2003b). These works offer by far the most comprehensive naturalistic theory of the first-person perspective that I know of.

I want to use Metzinger's view of the first-person perspective as a case study for naturalism. First, I'll present my own view of the first-person perspective (and point out its naturalistic and nonnaturalistic aspects), then I'll present Metzinger's reductive naturalistic account. After challenging some aspects of Metzinger's account, I want to consider some of the consequences of his account. Finally, I'll discuss naturalism more broadly and ask: Could there be a well-confirmed naturalistic theory that is rationally untenable and/or self-defeating?

1. THE FIRST-PERSON PERSPECTIVE: BAKER'S VIEW

All conscious beings—dogs, as well as human beings—have a perspective. They have points of view from which they perceive and act in the world. They solve problems by employing perspectival attitudes. Although a dog has a certain perspective on its surroundings with itself as "origin", the dog does not conceive of itself as a subject of experience. Metzinger puts it well:

> As Baker points out, it is not only necessary to have thoughts that can be expressed using 'I'. What is necessary is the possession of a concept of oneself as the *thinker* of these thoughts, as the *owner* of a subjective point of view. In short, what is needed is not only reference from the first-person point of view, but the capacity to mentally 'ascribe' this act of reference *to* oneself while it is taking place. (Metzinger 2003b, 396)

We not only *make* first-person references—e.g., 'I am registered to vote', but we also *attribute* first-person references to ourselves—e.g., 'I believe that I am registered to vote'. A first-person perspective[1] is a *conceptual* capacity to attribute first-person references to ourselves. This is a capacity to form complex first-person thoughts that have first-person references embedded in clauses following linguistic or psychological verbs. Call such thoughts and the sentences expressing them 'I*-thoughts' and 'I*-sentences', respectively. For example, 'I think (hope, fear, said) that I* am

---

[1] Throughout this paper, 'first-person perspective' should be understood as what I have lately called a 'robust first-person perspective' to distinguish it from a 'rudimentary first-person perspective' (Baker 2005).

tall' is an I*-sentence.[2] Note that I* thoughts include but are not limited to "Cartesian" thoughts (like 'I am certain that I* exist'). Mundane thoughts (like 'I hope that I* won't be late' or 'I wish that I* could buy a car') are I*-thoughts as well. I*-thoughts are first-person attributions of first-person reference, whereby one thinks of oneself as oneself*, without identifying oneself by means of any third-person referential device, such as a name, description, or demonstrative. Ability to express one's thoughts by means of I*-sentences is conclusive evidence of a first-person perspective.

From a first-person perspective, I have the ability to think of myself in a unique way, but there is no funny object that is myself-as-myself; there is no entity other than the person who I am. The referent of 'I' and of 'I*' is the person: not a body, not a disembodied ego. When I say, "I wonder whether I'll be happy in five years", I refer twice to myself—to the person, Lynne Baker, in her embodied concreteness. When I attribute first-person reference to myself by means of 'I*', what I <u>refer</u> to is no different from what you refer to by means of 'Lynne Baker'. What is special about my use of 'I*' is that I can conceive of that person in a way that you cannot, from "the inside," so to speak. This ability opens up a whole new realm of inwardness, of self-consciousness, of subjectivity.

On my view, having a first-person perspective is the defining characteristic of persons. What distinguishes us persons from other beings is our capacity to think of ourselves in a certain, first-personal way. A first-person perspective concerns *how* we think about ourselves, regardless of *what* we think about ourselves. If I am right, it is essential to your existence, to your being an entity in the world, that you have a first-person perspective. If you irretrievably lost your ability to think of yourself as yourself*, you would go out of existence—even if your brainstem still maintained the organic functions of your body. Since you the person and your body have different persistence conditions (yours depends on a first-person perspective; your body's depends on organic functioning), you are not identical to your body. On the other hand, you (the person) are not your

---

[2] Hector-Neri Castañeda introduced 'he*, and Gareth B. Matthews extended the he* from sentences with a third-person subject to 'I*' for sentences with a first-person subject. Castañeda studied phenomena expressed by sentences like 'The editor believes that he* is F.' See Castañeda 1966, and Castañeda 1967. Matthews discussed phenomena expressed by 'I think that I* am F'. See Matthews 1992.

body plus something else, just as a statue is not a piece of marble plus something else. The relation between you and your body (and the relation between the statue and the piece of marble) is what I've called 'constitution', a relation of unity that is not identity. (I worked out this view in *Persons and Bodies*, and in subsequent publications (Baker 2000, Baker 2002a, Baker 2002b).

The aim of my view of persons is to combine a fully Darwinian account of human organisms with a traditional concern of philosophers—namely, a concern with understanding our inwardness made possible by the first-person perspective. My account of the first-person perspective has some naturalistic and some non-naturalistic aspects. It is naturalistic in that it does not appeal to immaterial souls. The first-person perspective may well have evolved by means of natural selection; we human persons, with our first-person perspectives, are as much a part of the natural world as were dinosaurs.

I have no doubt that there's something going on in my brain that makes it possible for me to have I* thoughts, and I have no doubt that our capacity to have I* thoughts is a product of natural selection. The sub-personal sciences (e.g., neuroscience and parts of psychology) are sources of knowledge about mechanisms necessary for a first-person perspective in beings like us. But while I agree that the sciences may enhance our understanding of the *mechanisms* underlying the first-person perspective, I strongly disagree that knowledge of mechanisms can supplant or replace knowledge of phenomena that the mechanisms make possible.

Indeed, in many cases, knowledge of underlying mechanisms—though interesting in their own right—would not explain the phenomena that we set out to explain. For example, if we are interested in how winning the lottery changes the lives of lottery winners, a non-intentional explanation in terms of the natural sciences would be beside the point. And even where knowledge of underlying mechanisms is useful (as in, say, knowing the molecular events that trigger Alzheimer's disease), such knowledge does not unseat or replace knowledge of the disease as the destroyer of a person's life. In any event, I do not think that the natural sciences can

explain everything that there is to understand.[3] So, in this respect, I am not a naturalist.

Moreover, my Constitution view of persons may be considered to be non-naturalistic in other respects: One is that I deny that the biological origin of the first-person perspective tells us ontological significance of the first-person perspective. Ontology does not recapitulate biology. Now let us turn to Metzinger's view.

2. COGNITIVE CONSCIOUS SELF-REFERENCE: METZINGER'S VIEW

Metzinger writes sympathetically about my account of the first-person perspective. He writes that the conceptual distinction between merely having a perspective and conceiving of oneself as having a perspective—a distinction at the heart of my account of the first-person perspective—"is important for cognitive science in general, and also for the philosophical notion of a true cognitive subject" (Metzinger 2003b, 396). However, when I say, "[A]ttribution of first-person reference to one's self seems to be ineliminable" (Baker 1998, 331), Metzinger disagrees. He offers an alternative view that eliminates reference to any self or genuine subject of experience. On his view, "all that exists are conscious systems operating under transparent self-models" (Metzinger 2003b, 397). On my view, I (me, the person, a first-personal being, a genuine subject of experience, a "self") am an entity in the world. So, the issue between Metzinger and me is joined in a profound and intriguing way: When I affirm that there are persons with irreducible first-person perspectives in the world, I am affirming that there are genuine subjects of experience (essentially first-personal beings) in the world. When Metzinger denies that there are

---

[3] Some naturalists (e.g., Quine 1960) confine science to the so-called natural sciences; intentional descriptions are simply a dramatic idiom. I'll call this version Reductive Natualism. Other naturalists (e.g., Kornblith 1993) who are antireductionists may countenance irreducible social and psychological sciences that advert to intentional phenomena. Metzinger clearly aims for an account in terms of sub-personal mechanisms and is a reductive naturalist.

"selves," he is denying that there are genuine subjects of experience in the world.[4]

Let me make two terminological points:

(1) I follow Metzinger's use of the word 'phenomenal' to apply to the qualitative contents of conscious experience; phenomenal experience is characterized by how it feels or "what it's like" to have it. This leaves it open whether or not a phenomenal content represents anything real, or is, as Metzinger puts it, "epistemically justified" (Metzinger 2003b, 401). Phenomenal content may or may not depict anything in reality.

(2) Metzinger denies that there are any entities in the world that are "selves" or genuine subjects of experience. By the term 'genuine subject of experience,' I mean an entity that must be included as such in ontology—a first-personal entity that exists in the world and not just as an artifact of an information-processing system. Although I do not believe that there exist "selves" as distinct from persons, I do believe that there are persons, who are essentially first-personal, and are genuine subjects of experience (call them 'selves' if you'd like). I prefer the word 'persons' or 'genuine subjects of experience' to the word 'self', but I'll use all of these locutions to mean the same thing.

Although Metzinger emphasizes the importance of the first-person perspective in the very terms in which I describe it, he argues that we can account for the first-person perspective without supposing that there are "selves" or genuine subjects of experience. The question, then, comes down to this: Can there be an adequate ontology—an inventory of what really exists—that includes no first-personal subjects of experience, but only information-processing systems and self-models that are understandable in wholly third-personal terms?

The portion of Metzinger's argument that concerns me here has three parts:

---

[4] If all Metzinger means by a self or a subject of experience is "an internal and nonphysical object," Metzinger, 2003b, 271, then almost everyone agrees with him that there are none; and there would be no argument. I do not suppose him to be taking on a "straw man."

   (i) a sub-personal, naturalistic account of subjective experience,

  (ii) an account of how it can *seem* to us that we are genuine subjects of experience, and

(iii) an account of the (putative) fact that there really are no genuine subjects of experience in the world. Metzinger offers a theory both that denies that I am a genuine subject of experience and that shows what is really going on when it *seems* to me that I am a genuine subject of experience.

The first part of Metzinger's argument is to give an account of subjective experience. Our brains activate mental models that contain mental representations. Mental representations have both phenomenal content (smells, colors, etc.) that supervenes on brain states, and intentional content (wishing you were here, believing that global warming is a serious threat) that depend in part on relations to an environment. Our representations are part of mental models, some of which represent the world (world-models) and some of which represent the system generating the models (self-models).

> [A] self-model is a model *of* the very representational system that is currently activating it within itself. (Metzinger 2003b, 302) The content of a phenomenal self-model (PSM) is the conscious self: your bodily sensations, your present emotional situation, plus all the contents of your phenomenally experienced cognitive processes. (Metzinger 2003b, 299)

Some properties of a self-model are transparent—that is, we don't see them, we look through them; they are not introspectively accessible. Transparency here is a phenomenological, not an epistemological, notion. Other properties are opaque—that is, we are aware of them; they are introspectively accessible. E.g., as G.E. Moore pointed out, when we try to introspect the sensation of blue, the *sensation* (what the sensation of blue has in common with the sensation of green) is transparent: "we look through it and see nothing but the blue" (Moore 1903, 446). But the blue is opaque; it is what we see. Metzinger says:

> A transparent representation is characterized by the fact that the only properties accessible to introspective attention are their content properties. (Metzinger 2003b, 387)

Our subjective experience, in the first instance, is activation of representations in transparent models—i.e., only the representational contents are experienced, not the models themselves. In other words, subjective experience is phenomenal experience. It consists of activation of models of representations. We cannot experience the models. We experience only the content properties of representations, whether the contents depict anything outside the model or not.

The second part of Metzinger's argument is to show how it can seem to us that we are subjects of experience. Metzinger distinguishes between a phenomenal first-person perspective and a cognitive first person perspective (Metzinger 2003b, 405). A phenomenal first-person perspective allows an information-processing system to have phenomenal (i.e., subjective) experience; a cognitive first-person perspective allows an information-processing system to have I* thoughts that make it seem that it is a genuine subject of experience in the world.[5]

I*-thoughts require integrating part of an opaque self-model into a preexisting transparent self-model[6] (Metzinger 2003b, 402). The opaque self-model is a phenomenal model of the intentionality relation (PMIR) that "represents itself in an ongoing, episodic *subject-object relation*" (Metzinger 2003b, 411). What we think about when we consciously think about ourselves is really just the content of a self-model. In having I* thoughts, we are unable to consciously experience that "we are referring to

---

[5] To show how it can seem to us that we are subjects of experience, Metzinger begins with a transparent phenomenal self-model that can be generated by an animal or pre-linguistic being; then, a conscious cognitive subject emerges when the system generates opaque representations of itself and integrates them into the transparent phenomenal self-model. In Metzinger's words: "My claim is that, all other constraints for perspectival phenomenality satisfied, a conscious cognitive subject is generated as soon as a globally available representation of the system as currently generating and operating with the help of quasi-linguistic, opaque mental representations is integrated into the already existing transparent self-model" (Metzinger 2003a, 367f.; Metzinger 2003b, 395).

[6] Metzinger defines a minimal notion of self-consciousness as having three properties: "the content of the self-model has to be embedded into a currently active world-model; it has to be activated within a virtual window of presence; and it has to be transparent" (Metzinger 2003a, 373).

the content of a *representation* that is 'in ourselves' (in terms of locally supervening on brain properties)" (Metzinger 2003b, 402).
Metzinger continues:

> Cognitive self-reference always is reference to the phenomenal content of a transparent self-model. More precisely, it is a *second-order* variant of phenomenal self-modeling, which, however, is mediated by *one and the same* integrated vehicle of representation. The capacity to conceive of oneself as oneself* consists in being able to activate a dynamic, 'hybrid' self-model: Phenomenally opaque, quasi-symbolic, and second-order representations of a preexisting phenomenally transparent self-model are being activated and continuously reembedded in it. This process is the process of [conscious cognitive self-reference]....Reflexive self-consciousness consists in establishing a subject-object relation within the [phenomenal self-model]. [7] (Metzinger 2003b, 403)

Let me try to put this in my own words. If someone thinks, "I am hungry," she is activating a transparent phenomenal self-model. She sees through the 'I' (so to speak) to the feeling of hunger. The 'I' is invisible to her. But if she thinks, "I believe that I* am hungry," the first occurrence of 'I' is part of an opaque self-model that is integrated into the preexisting transparent self-model. The second occurrence of 'I' in 'I believe that I* am hungry' (the 'I*') is phenomenologically transparent. The first occurrence of 'I' is opaque since she is thinking of herself as the subject of her thought. What remains invisible to her is precisely what she is referring to. A conscious information-processing system seems to be a subject of experience when it generates subjective experiences that include the experience of being a subject of experience. Thus, we seem to be subjects of experience in the world. But the experience of being a subject of experience remains phenomenal.

The third part of Metzinger's argument is to show that the experience of being a substantial subject is *merely* phenomenal. The conscious cognitive subject is not part of reality, but only part of a self-model. Metzinger holds that a cognitive first-person perspective (that is, the ability to have I*

---

[7] Metzinger 2003b, 403. I inserted 'consciously experienced cognitive self-reference' for 'introspection$_4$'. Metzinger characterizes introspection$_4$ as "a conceptual (or quasi-conceptual) kind of metarepresentation, operating on a pre-existing , coherent self-model" (Metzinger 2000a, 367).

thoughts) is a special case of a phenomenal first-person perspective: "Cognitive self-reference is a process of phenomenally modeling certain aspects of the content of a preexisting transparent self-model, which in turn can be interpreted as the capacity of conceiving of oneself as oneself*" (Metzinger 2003b, 405). In cognitive self-reference, what is referred to is the phenomenal content of a transparent self-model. So, the reference will be to an element of the self-model, not to a self existing in the world. In short, the conscious cognitive subject is just an element of the self-model. Metzinger says:

> Any conscious system operating under a phenomenally transparent self-model will by necessity instantiate the phenomenal property of selfhood in a way that is untranscendable for this system itself. (Metzinger 2003a, 363)

I believe that the word 'untranscendable' in this passage means that the system lacks resources to uncover the fact that the phenomenal property of selfhood is *merely* the content of a self-model. But according to Metzinger, what we refer to in cognitive self-reference is a mental representation: "[I*]," he says, "is the content of the transparent self-model. (Metzinger 2003b, 400).

Metzinger's claim that the *cognitive* first-person perspective can be reduced to a complex *phenomenal* first-person perspective has a strong consequence about subjects of experience: No belief about the worldly existence of *what* is being mentally represented is "epistemically justified." That is, we cannot conclude that what is represented exists in reality. Metzinger says that the belief that a self carries out the act of cognitive self-reference is not epistemically justified, and hence is apt for rejection (Metzinger 2003b, 403). Thus, we can see how the Cartesian claim of epistemic transparency (my certainty that I am a genuine subject of experience that exists in reality,) is intelligible, even if it is false (Metzinger 2003a, 363).

In sum, Metzinger denies that conscious experience really has a subject in the world (a self or person who does the experiencing). Our experience of being subjects of experience is only phenomenal. We are mistaken if we think that, because we experience being a subject of experience, there actually *is* (in reality) a subject of experience who we are. We lack "epistemic justification" for "all corresponding belief states about what is

actually being represented" (Metzinger 2003b, 404; Metzinger 2003a, 375). The subjective experience of being someone in the world is an illusion. Just as dreams and hallucinations tell us nothing veridical about what's really going on in the environment, so too does subjective experience tell us nothing veridical about what we are. There are no selves, just self-models. "For ontological purposes," he says, "'self' can therefore be substituted by 'PSM' [phenomenal self-model]" (Metzinger 2003b, 626).

Metzinger says that the main thesis of his book *Being No One*

> is that no such things as selves exist in the world: Nobody ever was or had a self. All that ever existed were conscious self-models that could not be recognized as models. (Metzinger 2003b, 1)

The experience of oneself is only a phenomenological consequence of a system operating under a phenomenal self-model (Metzinger 2003b, 387). This is compatible with saying either that I (a subject of experience) do not exist, or that I exist but that what I am is only a part of the content of a self-model.

However, I believe that the most charitable way to read Metzinger is not as an eliminativist about subjects of experience, but as a reductionist. Despite the misleading title of his book *Being No One*, and despite what I just quoted him as saying, perhaps he is not saying that I do not exist, or that I am no one. Perhaps he is saying that *what I am* is an information-processing system that has generated a phenomenal self-model (PSM), and that *what I think about* when I think about myself is only the content of a mental representation in my self-model.

In any case, whether Metzinger is an eliminativist about selves (as his quotations suggest) or a reductionist (as I think is the more charitable interpretation), he denies that there exist what I have called 'genuine subjects of experience'—first-personal entities that must be included as such in ontology. If Metzinger is correct, then the fact that you and I seem to be subjects of experience has no ontological significance. Persons (selves, subjects of I* thoughts) belong to appearance, not to reality.

3. TWO ISSUES INTERNAL TO METZINGER'S THEORY

Let me express my admiration for the cleverness of Metzinger's theory. Indeed, there are a number of points of broad agreement between Metzinger and me. Here are some examples:

(1) Self-consciousness is importantly different from mere sentience, or the kind of consciousness that nonhuman animals have (Metzinger 2003b, 396).

(2) Self-conscious beings possess the distinction between the first and third person "on a *conceptual* level, and actually use it" (Metzinger 2003b, 396).

(3) Philosophers cannot "decide on the truth or falsity of empirical statements by logical argument alone" (Metzinger 2003b, 3).

(4) The phenomenology of conscious experience should be taken seriously (Metzinger 2003b, 301 n2).

(5) A human being can "conceive of itself *as a whole*" (Metzinger 2003b, 1).

Despite these areas of agreement, I would like to critically discuss two issues internal to Metzinger's view, and then turn to the main difference between my view and Metzinger's: The main difference between us is the ontological difference, stemming from his commitment to reductive naturalism. Whereas I think that a complete ontology must include persons ("selves" or genuine subjects of experience), Metzinger does not. That is, although I think that there are selves in reality (again, I really prefer the word 'person'), Metzinger thinks that selves are only matters of appearance, not reality. On his view, as we have seen, reality includes no selves, only self-models.

    The two issues internal to Metzinger's view that I want to discuss are, first, Metzinger's "analysis" of cognitive first-person reference from a third-person point of view, and second, his notion of phenomenal content and the use that he makes of it.

    First, consider Metzinger's argument against my claim that attribution of first-person reference to oneself is ineliminable. In the article of mine

that Metzinger discusses, I used the example of Descartes' I*-thought, [I am certain that I* exist][8], and I pointed out that the certainty that Descartes claimed was first-personal: Descartes claimed that he was certain that he* (he himself) existed, not that he was certain that Descartes existed. Although Metzinger agrees that Descartes was not making a third-person reference to Descartes (Metzinger 2003b, 398), he also holds that the mental content of Descartes' thought [I am certain that I* exist] and the linguistic content of the sentence 'I am certain that I* exist' can be understood in third-person terms.

All the mental content of the thought [I am certain that I* exist] is merely phenomenal and, as Metzinger says, "not epistemically justified" (Metzinger 2003a, 373).[9] In short, my certainty that I* exist is understood as a complex relation of parts of the content of a self-model. In general, I*-thoughts are to be understood without supposing that a subject of experience exists in reality.

Metzinger also treats *linguistic* self-reference by the sentence <I am certain that I* exist>. The linguistic content of <I am certain that I* exist> may be "analyzed", he says, from a third person perspective as follows:

> (A) <The speaker of this sentence currently activates a PSM (a phenomenal self-model) in which second-order, opaque self-representations have been embedded. These representations are characterized by three properties:
>
> First, they possess a quasi-conceptual format (e.g., through a connectionist emulation of constituent-structure, etc.);
>
> second, their content is exclusively formed by operations on the transparent partitions of the currently active PSM;
>
> third, the resulting relation between the system as a whole and content is phenomenally modeled as a relation of certainty.> (Metzinger 2003b, 402)

---

[8] Metzinger uses square brackets ([....]) to denote thoughts, and pointed brackets (<....>) to denote linguistic expressions.

[9] Such phenomenal certainty has two defining characteristics. The first is that "the object-component of the phenomenal first-person perspective is transparent and the respective person is therefore, on the level of phenomenal experience, forced into an (epistemically unjustified) existence assumption with respect to the object-component." The second defining characteristic is "transparency of the self-model yielding a phenomenal self depicted as being certain" (Metzinger 2003a, 374).

Let us label this account (A). Can (A) be a correct analysis of a first-person assertion <I am certain that I* exist>? My assertion <I am certain that I* exist> is necessarily about me, Lynne Baker. But the analysis is not. The analysis is about anybody who asserts that she* is certain that she* exists. Neither my assertion <I am certain that I* exist> nor (A) entails the other. So, the proposed analysis (A) is not an analysis in a traditional sense. Nor can (A) replace anyone's assertion of 'I am certain that I* exist.' The target sentence and (A) simply do not convey the same information.[10]

What is at issue is not the specific Cartesian example <I am certain that I* exist>, however, but rather my broader claim that the attribution of first-person reference to one's self seems to be ineliminable," (Baker 1998, 331). It is this broader claim—one that applies to all I*-thoughts and I*-sentences that is at stake.

So perhaps (A)—even if it is not an analysis—should be regarded as an application of part of an empirical theory. Metzinger predicts that the phenomenal self-model (PSM) is a real entity that will be empirically discovered—"for instance, as a specific stage of the global neural dynamics in the human brain, characterized by a discrete and unitary functional role" (Metzinger 2003b, 411). The only thing to say here is that we will have to wait and see whether neural correlates of phenomenal self-models are actually discovered in the brain.

Even if they are discovered, however, the most that a third-person empirical theory of I*-sentences or I* thoughts can hope to do is to provide necessary and sufficient conditions for the production of I* sentences or I* thoughts. But this would be a far cry from eliminating or replacing I* sentences or I* thoughts by third-person sentences or thoughts. Even if (A) is part of an empirical theory that is eventually confirmed, it still cannot *replace* the I*-sentence, which remains ineliminable.

The second question that I want to raise that is internal to Metzinger's theory is whether the notion of phenomenal content can bear the load that Metzinger puts upon it. Phenomenal content is qualitative content and

---

[10] Maybe (A) is what makes an assertion of <I am certain that I* exist> true. Maybe (A) is the truth-masker for such assertions. But the notion of truth-makers is part of a controversial metaphysical theory outside the purview of any empirical science known to me. So, as a naturalist, Metzinger should be reluctant to appeal to truth-makers. (And, as far as I know, he does not appeal to truth-makers.)

(supposedly) supervenes on the brain; representational content is intentional content (Metzinger 2003b, 71). Metzinger says:

> The central characteristic feature in individuating mental states is their phenomenal content: the way in which they *feel* from a first-person perspective. (Metzinger 2003b, 71)

In my opinion, this is not the way that mental states should (or even could) be individuated—at least those mental states that have truth-conditions, as all I*-thoughts do. We have no criterion for sameness of feeling: I wake up at night and on some occasion my subjective experience is hope that I'll get a certain paper finished on time; on another occasion, my subjective experience is hope that it won't rain tomorrow. My subjective experience is certainly not the same on both occasions of hope, but not because of any difference in feeling. The difference—even the difference in what it's like to be in the states—depends on the *intentional content* of the hopes, not on any feeling associated with them. So, I do not think that purely phenomenal content can individuate mental states.

  According to Metzinger, "conceptual forms of self-knowledge" (I* thoughts) are generated

> by directing cognitive processes towards certain aspects of internal system states, the intentional content of which is being constituted by a part of the world depicted as *internal*. (Metzinger 2003a, 367; his emphasis.)

He says that the phenomenology associated with this type of representational activity

> includes all situations in which we consciously think about ourselves *as* ourselves (i.e., when we think what some philosophers call I* thoughts; for an example see Baker 1998). (Metzinger 2003a, 367)

It seems to me to be phenomenologically mistaken to suppose that the intentional contents of I* thoughts depict part of the world as internal. When I think: "I believe that I* can get money from this ATM", the intentional content of my I*-thought is not constituted by a part of the world depicted as *internal*. Still less is internality "phenomenally experienced." When I consciously think, "I believe that I* can get money

from this ATM," the intentional content of my thought depicts a relation between a machine and myself—a relation that is not internal to me.

Metzinger endorses a principle of local supervenience for phenomenal content: "phenomenal content supervenes on spatially and temporally internal system properties" (Metzinger 2003b, 112). He goes on:

> If all properties of my central nervous system are fixed, the contents of my subjective experience are fixed as well. What in many cases, of course, is not fixed is the *intentional* content of those subjective states. (Metzinger 2003b, 112)

But almost all subjective experience (mine, anyway) has intentional content. Any mental state that can be true or false, or that can be fulfilled or unfulfilled, has intentional content, no matter what it feels like.[11]

For example, it suddenly occurs to me that I locked my keys in my office, and I experience a feeling of panic. The subjective experience has intentional, not just phenomenal, content; it includes a thought that has a truth value. And I'm greatly relieved if I discover that the truth value of my thought is false: Here the keys are in my pocket. The subjective experiences were not just the panic and the relief; they included the sudden thought with its specific intentional content and the happy discovery that the thought was false. Not only are we embodied, but also we are embedded—embedded in a real world, not just in representations of a world. And the contents of our subjective experience are typically infected by relations with the environment.[12]

Since, according to Metzinger, phenomenal content supervenes on brains, and most of our subjective experience has intentional content, which does not supervene on the brain, phenomenal content cannot account for our subjective experience. Our brains, and what supervene on them, are only one determinant of subjective experience. I may wake up in

---

[11] Although I cannot argue for it here, I believe that none (or almost none?) of our intentional mental states supervene on our brain states. See, Baker, forthcoming.

[12] Metzinger notes that one "of the most important theoretical problems today consists in putting the concepts of 'phenomenal content' and 'intentional content' into the right kind of logical relation" (Metzinger 2003b, 112). That seems to me a problem easily solved: Do not insist that phenomenal content (content that is experienced) supervene on brain states. With the exception of qualia (if there are any), all content depends on interaction with the environment.

the night, thinking that a search committee meeting the next day may be unpleasant. That particular subjective experience would be metaphysically impossible (and not just causally impossible) in a world without search committees and all the intentional apparatus surrounding hiring new people. So my subjective experience of thinking that tomorrow's meeting may be unpleasant does not supervene on my brain. Hence, phenomenal content, which does supervene on my brain, does not suffice for ordinary subjective experience. Metzinger asserts:

> Phenomenal content can be dissociated from intentional content: a brain in a vat could possess states subjectively representing object colors as immediately and directly given. (Metzinger 2003a, 359)

This claim brings to the fore the dilemma that phenomenal content faces: If phenomenal content is dissociated from intentional content, it does not account for much of our subjective experience, as the above examples show. But if phenomenal content is not dissociated from intentional content, then phenomenal content does not supervene locally on brain states and it loses the neuroscientific legitimacy that Metzinger claims for it. Either way, phenomenal content cannot play the role that Metzinger assigns it.

To recapitulate, my two objections internal to Metzinger's view concern his attempt to eliminate I*-thoughts and I*-sentences (or to reduce them to the third-person), and his use of phenomenal content to carry the weight of subjective experience. Now let us turn to some consequences of Metzinger's theory.

4. CONSEQUENCES OF METZINGER'S THEORY

In this section, I want to consider three kinds of consequences of Metzinger's view that I find untenable—semantic, epistemic, and moral consequences.

First, I believe that Metzinger's view requires an ineliminable equivocation on the word 'I'. Sometimes 'I' refers to the whole information-processing system, and sometimes 'I' refers to the content of a part of a self-model. This becomes apparent if we consider I* sentences.

Consider an ordinary I* thought—e.g., 'I believe that I am in Austria'. Metzinger says:

> I experience myself as the thinker of the I*-thoughts. (Metzinger 2003a, 373)

The reality that the first occurrence of 'I' in this thought refers to is the whole information-processing system.

> The content of [I] is the thinker, currently representing herself as operating with mental representations. (Metzinger 2003b, 401)

It is the whole system that thinks of itself as the thinker of thoughts. On the other hand, the second occurrence of 'I' (the 'I*' in 'I believe that I am in Austria') "is the content of the transparent self-model." As Metzinger explains:

> Any conscious system operating under a transparent self-model will by necessity instantiate a phenomenal self to which, linguistically, it *must* refer using <I*>. (Metzinger 2003b, 400, emphasis his.)

So, the referent of 'I' is sometimes the whole information-processing system and sometimes the content of a self-model. It is utterly implausible that 'I' could be equivocal in a single thought of a single thinker. This would make us all hopelessly schizophrenic: Which am I—the whole information-processing system or part of the transparent content of its currently active self-model?

   We can see this tension in another way when we consider Metzinger's metaphor that

> you constantly confuse yourself with the content of the self-model currently activated by your brain. (Metzinger 2003b, 1)

Who is doing the confusing? On the last page of his book, Metzinger says that we should not take this metaphor too literally:

> There is no one *whose* illusion the conscious self could be, no one *who* is confusing herself with anything. (Metzinger 2003b, 634).

What, exactly, then is the confusion that has no bearer? It is difficult to see how there is a confusion to be made (with or without someone to make it).

When I think, "I believe that I* in Austria", my belief is that I (all of me) am in Austria. Perhaps, Metzinger is saying that, unbeknownst to me, the information-processing system that I am has a transparent self-model representing being in Austria, and the system integrates part of an opaque self-model representing itself into the transparent self-model, and thus generates a representation of a representation of being in Austria within the self-model. This would completely misrepresent the content of my thought "I believe that I* am in Austria." If you and I agree that I believe that I* am in Austria, then we are agreeing about me, about where I believe I am (even if I am an information-processing system operating with a self-model); we are not agreeing about my self-model. So, I think that it is not coherent to construe the subjects of I* thoughts to be parts of self-models.

Second, consider an epistemic consequence of Metzinger's view. The theory cannot make sense of what is going on when people reflect on what they are doing while they are doing it. Suppose that a scientist using an electron microscope for the first time thinks to herself, "I can hardly believe that I'm looking at electrons." If the scientist is not a subject of experience that exists in the world, how is she to make sense of her own thought, on Metzinger's view? Well, maybe this: The scientist has the experience of being the subject of the thought expressed by "I can hardly believe that I'm looking at electrons," but she is not "epistemically justified" in supposing that she really is a genuine subject of experience in the world.

From Metzinger's point of view, the scientist is an information-processing system that is integrating

> its own operations with opaque mental representations, i.e., with mental simulations of propositional structures that could be true or false, into its already existing transparent self-model while simultaneously attributing the causal role of generating these representational states to itself. (Metzinger 2003a, 369)

But, on Metzinger's view, *the scientist herself* cannot see her own thoughts and activity in this light; indeed, she is deceived about what is going on. Of course, Metzinger has an account of why she cannot see her own thoughts and activity in this light; but that's beside the point. The point is that the scientist cannot comprehend what is really going on while she is engaging in scientific activity. Metzinger's theory would seem to make it impossible

for anyone to think clearly about what she is doing while she is doing it. A view of subjectivity that makes it impossible for scientists (and everyone else) to think clearly about what they are doing as they are doing it is dubious.

Third, Metzinger's view has consequences that are morally questionable. Consider a soldier long ago who experienced excruciating pain while undergoing a battlefield amputation. Metzinger says that we should minimize "the overall amount of suffering in all beings capable of conscious suffering" (Metzinger 2003b, 570). I do not see what epistemic grounds we can have for this "simple principle of solidarity," as he calls it. If Metzinger's view is correct, then we are epistemically unjustified in supposing that there is any substantial entity in the world that actually undergoes excruciating pain; rather, there is an information-processing system with a self-model that made it appear that there was such a subject of pain. There was a subjective experience of pain, but the bearer of the pain was just a phenomenal self, who was "epistemically unjustified." If we are unjustified in supposing that there was a substantial entity (the soldier) who was a subject of pain, then we would be under no obligation to alleviate the pain. I think that this consequence would make our moral experience unintelligible.[13]

I am prepared to accept theories with counterintuitive consequences (e.g., I find it counterintuitive that there's no absolute ongoing now; but I accept this as a result of well-confirmed theories of physics). But Metzinger's view of the first-person perspective and its I* thoughts is not just counterintuitive. It has consequences that seem to me to be semantically, epistemically and morally untenable. So, what should we do?

5. WHITHER NATURALISM?

Metzinger's theory is a naturalistic one. Naturalism is often characterized by two themes—an ontological one that is committed to an exclusively

---

[13] In an email to me, Metzinger said that he was very interested in ethical consequences of his view. He said that he believes that there can be selfless suffering subjects, and that phenomenal suffering is real and should be minimized. I hope that he pursues these issues at length. It is not obvious to me how to work out a morally acceptable position within the confines of his view.

scientific conception of nature, and a methodological one that conceives of philosophical inquiry as continuous with science (De Caro & Macarthur 2004, 3). Reductive naturalism recognizes as real only third-personal entities and properties.[14]

Metzinger's third-person sub-personal account of the first-person perspective fits this characterization of reductive naturalism nicely. So, I shall continue to use Metzinger as a case study. On being presented with a theory, each of us decides: Do I accept this theory? I invite you to join me in thinking of Metzinger's theory from the point of view of a prospective adherent of it. Would it be *rational* for me to accept it? Would it even be *possible* for me to accept it? Let's consider each of these questions in turn.

(i) If Metzinger's view is correct, then there are no selves and no genuine subjects of experience in the world. I just argued that without subjects of experience in reality, I cannot make sense of my own experience while I'm having it. A view with this consequence renders my experience unintelligible to me. Is it rational for me to endorse a theory that renders my experience unintelligible to me? My experience of being a conscious subject is evidence that I am a subject, and this evidence overwhelms any possible evidence that I may have for any scientific theory to the contrary. Hence, rationally, I should reject the view that would have me repudiate myself as a genuine subject of experience.

(ii) It seems that Metzinger's theory cannot coherently be endorsed or accepted. I may have the subjective experience that I* am accepting Metzinger's theory. I think to myself, "I am having the experience that I* am accepting Metzinger's theory." But the "I*" doing the accepting is not an entity in the world; it is just part of the content of a transparent self-model (Metzinger 2003a, 372; Metzinger 2003b, 400). When I refer to myself by means of 'I*', I am referring to the content of a mental representation. It is incoherent to suppose that a mental representation can actually accept a theory. On Metzinger's view, all there can be is a subjectless subjective experience of accepting his theory; but for me to accept a theory is not just for there to be a subjectless subjective experience of accepting. So, it seems doubtful that Metzinger's theory can be endorsed

---

[14] Whether non-reductive naturalism can allow irreducibly first-person phenomena remains to be seen.

or accepted. If a theory cannot coherently be endorsed or accepted, it is self-defeating. It is paradoxical, if not self-contradictory, to suppose that I should accept a theory that I cannot coherently accept.

Here is my recommendation: Give up reductive naturalism. Do not confine ontological conclusions to those that can be gleaned by scientific methods. As we have seen in the best attempt to naturalize the first-person perspective, science (at least as it stands today) cannot intelligibly be the final word on what there is. Even if philosophers gave up naturalism as a *global* commitment to the methods and ontology of natural sciences, however, we may still keep those naturalistic theories that explain what we want explained. The way to accomplish this is to attend to what the naturalistic philosophical theories are (or should be) theories *of*.

We should distinguish between phenomena that interest philosophers and the underlying mechanisms that subserve those phenomena. For example, we may hope for a naturalistic theory of the mechanisms that underwrite a first-person perspective (Metzinger 2003b, 395). But on my view, the "I" who is the genuine subject of experience is a person: an object in the world whose first-person perspective is irreducible and ineliminable.

Why is my view to be preferred to Metzinger's? First, his theory (with a phenomenal self that is not a genuine object in reality) is paradoxical; mine is not. Second, his theory relies on an inadequate view of subjective experiences as supervening on the brain; mine does not. Third, his theory would leave the work undone that the first-person perspective does—e.g., in understanding moral agency; mine does not (Baker 2000). Fourth, his interpretation has unfortunate semantic, epistemic, and moral consequences; mine does not.

Reductive naturalism often seems like a change of subject that lacks respect for the peculiar projects and puzzles that traditionally preoccupy philosophers. In particular, nonnaturalists resist the tendency to assimilate the phenomena that piqued our philosophical interest to the mechanisms that support those phenomena. No one doubts that there are underlying mechanisms and that they are worthy of understanding. The nonnaturalist resistance is to *supplanting* philosophical questions by empirical questions about the underlying mechanisms that make the philosophically-interesting phenomena possible—as if questions about the 1985 world-championship

chess match between Kasparov and Karpov could be replaced by questions about the physics involved in the motions of little bits of wood.

Taking Metzinger's view as the best case, I now suspect that the challenge that the first-person perspective poses for reductive naturalism cannot be met.[*]

6. REFERENCES

Baker, L. R. 1998: "The First-Person Perspective: A Test for Naturalism." *American Philosophical Quarterly* 35, 327-348.

Baker, L. R. 2000: *Persons and Bodies: A Constitution View*. Cambridge: Cambridge University Press.

Baker, L. R. 2002a: "On Making Things Up: Constitution and its Critics." *Philosophical Topics* 30, 31-51.

Baker, L. R 2002b: "Precis" and "Reply to Critics." *Philosophy and Phenomenological Research* 64, 592-598 and 623-635.

Baker, L. R. forthcoming: "First-Person Externalism". *The Modern Schoolman.*

Block, N. 1995: "On a Confusion About a Function of Consciousness." *Behavioral and Brain Sciences* 18, 227-247.

Castañeda, H.-N., 1966: "He: A Study in the Logic of Self-Consciousness." *Ratio* 8, 130-157.

Castañada, H.-N.. 1967: "Indicators and Quasi-Indicators." *American Philosophical Quarterly* 4, 85-100.

Chalmers, D. J. 1996: *The Conscious Mind: Toward a Fundamental Theory*. Oxford: Oxford University Press.

De Caro, M. and Macarthur, D. 2004: *Naturalism in Question*. Cambridge, MA: Harvard University Press.

Kornblith, H. 1993: *Inductive Inference and its Natural Ground*: *An Essay in Naturalistic Epistemology*. Cambridge, MA: MIT Press.

Matthews, G. B. 1992: *Thought's Ego in Augustine and Descartes*. Ithaca. NY: Cornell University Press.

Metzinger, T. 2003a: "Phenomenal Transparency and Cognitive Self-Reference." *Phenomenology and the Cognitive Sciences* 2, 353-393.

Metzinger, T. 2003b: *Being No One: The Self-Model Theory of Subjectivity*. Cambridge, MA: MIT Press.

Moore, G. E. 1903: "The Refutation of Idealism." *Mind* 12, 433-53.

Quine, W. V. O. 1960: *Word and Object*. Cambridge, MA: MIT Press.

Which Ontology for Naturalists?

Josef Quitterer
University of Innsbruck

According to classical naturalism, folk psychology (our system of mental explanations) is compatible with scientific explanations of human behaviour. There are *proper physical correlates* for mental phenomena. Contemporary naturalistic approaches do not share this compatibilistic view. According to them, folk psychological explanations are incompatible with scientific explanations because they cannot be separated from specific ontological commitments which comprise, among other things, endurers. In the article I present the following thesis: Most contemporary *naturalists* are committed to the view that there are *no proper physical correlates* for mental phenomena because they allow only events and processes as physical correlates and not endurers or continuants. I argue against the naturalistic assumption that the commitment to event ontology follows from the assumption that reality can be explained in scientific terms. For this purpose I will concentrate on the naturalistic explanation of the phenomenon of the human self.

1. FROM COMPATIBILISM TO INCOMPATIBILISM

According to the classical naturalistic theory of reductionism there is no contradiction between our everyday folk psychological assumptions and scientific explanations. In this view, the explanatory success of our common sense or folk psychological assumptions is based upon an isomorphic structure of mental and physical events (Lewis 1966): We are justified to explain our actions by assuming mental causes like wants and beliefs because these mental causes correspond to neural patterns which are the physical causes of our actions. According to Goldman, Lewis, Fodor, and many others, there is no contradiction between folk psychological and neurobiological explanations of human behaviour. Cognitive science is not here to correct and revise our common sense

assumptions but to clarify the "characteristic manner in which wants and beliefs cause acts":

> […] neurophysiological information can help explain how it is that wants and beliefs cause action. (Goldmann 1970, 166-168)

This positive attitude towards our basic folk psychological assumptions is compatible with the naturalistic commitment that the only form of reality which exists is physical in nature. The theory behind this compatibilist view is the model of theory reduction. It is maintained that the folk psychological system of our mental concepts can be reduced upon the system of our best scientific theories. For classical naturalism scientific descriptions were much more adequate and closer to reality than mental descriptions. But that did not exclude the belief that also the mental concepts had a *fundamentum in re* because in the end they referred to the same (physical) reality as physical descriptions did. In this view mental states had a 'proper physical correlate'. '*Proper physical correlate*' is defined here as a *physical entity whose causal role corresponds to the causal role of the mental entity*.

The assumption of proper physical correlates of mental phenomena is not shared any more by most contemporary naturalists.[1] There are several reasons for this change of attitude within naturalism towards mental phenomena and folk psychology. Classical naturalism was focused on mental phenomena which had an event-like and relational character—mental events like wants, intentions, desires, pains etc. It seemed to be easy to find corresponding correlates on the physical level upon which these phenomena could be reduced. However, a closer analysis of mental phenomena revealed the difficulties of the reductionist enterprise of classical naturalism: It turned out that these event-like mental phenomena had a qualitative aspect which could not be grasped in the reduction. This

---

[1] The list of naturalists who deny the existence of proper physical correlates comprise philosophers and cognitive scientists like Dennett, Flanagan, Metzinger, Roth, Crick & Koch, Paul and Patricia Churchland, Stich, Wegner and many others. There are some exceptions, e. g. Kim and Beckermann; these naturalists can be regarded as the last representatives of classical naturalism.

difficulty leads David Chalmers to his well-known distinction between hard and easy problems of consciousness (Chalmers 1995).

The main problem was, though, that the descriptions of mental phenomena cannot be separated from a whole set of descriptions which together form the conceptual system of folk psychology (and folk ontology). "Folk psychology" is defined according to a mainstream position in the theory of mind-debate as

> a conceptual framework ... used by ordinary people to understand, explain and predict their own and other people's behaviour and mental states. (Von Eckardt 1994, 300)

Not only belong the above mentioned mental concepts like belief, desire etc. to the conceptual system of folk psychology but also a whole range of other concepts which are implicitly presupposed in folk psychological (mental) explanations: freedom of will, agent causation, the existence of persons etc.

The decisive point is that there are also specific ontological commitments which are connected with the conceptual system of folk psychology. One of the most important ontological commitments of folk psychology is the assumption of the existence of enduring subjects which are the bearers of intentional states. Propositional attitudes presuppose acting and thinking subjects which remain the same during time. Endurers are presupposed in nearly all commonsense explanations. Paul Churchland emphasizes that our common-sense framework for empirical reality contains a subframework comprehending the notion of a person (Churchland 1979, 89). We presuppose that the things to which we attribute mental states are persons which remain the same during their mental activity. Diachronic personal identity makes it possible that we can ascribe mental states also during these periods in which the subject does not have occurrent mental activity. Intentions, beliefs, and wants do not exist only as mental events or occurrences—"I now want to go skiing", "I am now aware of my belief that there is snow in the mountains"—but also in the dispositional form of so called 'standing' wants, desires or intentions: "Since a few weeks I have the intention to go skiing as soon as there will be enough snow", "Since my 10th birthday I have the desire to obtain the pilot's licence." These mental states exist even when they are

not present in an episodic form, even when we are not actually aware of them. Mental dispositions cannot be reduced upon underlying physical events or facts[2] and we ascribe dispositions to objects, not to properties or events. Obviously dispositions need for their existence a bearer, which is definitively not event-like, but a continuant or a substance.[3] Without enduring subjects the attribution of long-term mental dispositions would not make sense. Enduring subjects are also presupposed in our ascriptions of mental causation: An agent subject can be the intentional cause of her acts only if she remains the same at least until the intended action has come to its conclusion.

It is clear that in the everyday language of folk psychology we do not use technical terms like 'endurers', 'continuants' or 'substances' in order to express our intuition that we remain the same during time. When we justify our common sense belief that—despite all our physical and psychical changes—we are now the same persons as we were ten years ago, we use concepts like 'I' or 'self': "Yes, I looked totally different ten years from now, but it's still me, my self remained unchanged."

For most contemporary naturalists in the world of physical events there can be no *proper physical correlate* for the folk psychological assumptions of a self. The concepts of 'self' cannot be reduced upon something physical as it was attempted in the case of mental events by physicalists in the 'good old times' of classical reductionism. Contemporary naturalists assume that there is a physical correlate for the self in a similar way as there is one for illusions or false beliefs: Something must be going on in a person's brain when she is convinced that she is Napoleon when in fact she is not; in this sense, the belief "I am Napoleon" has a physical correlate. But there is no *proper physical correlate* for this illusion—there is no state of affairs in the physical world which corresponds to the propositional content of the belief that she is Napoleon. A similar opinion is shared by naturalists when it comes to folk psychological assumptions like self and personal identity: There is nothing in the physical world which corresponds in the proper sense to that what we have in mind when we refer to our selves. For

---

[2] For the problems concerning the reduction of dispositions see Mumford 1998.

[3] According to Jansen 2007, 168, disposition ascriptions to individual substances are more basic in metaphysical, logical and epistemological respect.

Metzinger, e.g., the self is the product of the self-misunderstanding of a system which self-represents itself (Metzinger 1993, 157), for Dennett the self is an explanatory fiction (Dennett 1991, ch. 13).

It is obvious that this new naturalistic approach has negative consequences for the relationship between science and folk psychology. Due to the ontological commitments of folk psychology it is not possible to reduce smoothly our everyday assumptions upon physical descriptions. The entire folk psychological system and its ontology are incompatible with scientific knowledge. Eliminative materialists deduce from that the necessity to eliminate the folk psychological framework in favour of a more sophisticated neurobiological one:

> Eliminative materialism is the thesis that our common-sense conception of psychological phenomena constitutes a radically false theory, a theory so fundamentally defective that both the principles and the ontology of that theory will eventually be displaced, rather than smoothly reduced, by completed neuroscience. (Churchland 1990, 206)

Eliminative materialism is an extreme position and therefore not widely accepted among naturalists. Even if folk psychology is based upon wrong propositions, most naturalists would admit that our folk psychological system is very useful in everyday life. For this reason, they stress the evolutionary advantage of such false assumptions concerning the self and diachronic personal identity. They tend to explain the self or related folk psychological assumptions as some sort of useful fictions (Dennett 1991, Flanagan 1992, Metzinger 1993, Roth 2003). According to these authors it is only our ability to self-represent ourselves that can be proved scientifically. This ability, however, does not presuppose a robust notion of the self. It is sufficient to talk about a biological system controlling and representing itself. According to this position the 'proper self' is nothing else than a fiction created by a self-representing organism. Physical (and thus scientifically observable) correlates of self-representation are neuronal activities taking place in the brain of the self-representing biological system (Dennett 1991, 187).

This naturalization of our assumption of diachronic identity expressed by concepts like 'self' or 'I' is officially justified through scientific evidence: The thesis that the self is an illusion or fiction is deduced, e.g. by

Dennett, from the fact that the physical correlates of the psychological self consist in parallel distributed neural processes in a highly plastic brain. For Dennett it is the specific complexity and plasticity of the human brain which makes it possible that the folk psychological self can detach itself completely from the biological organism. Cases of DID (Dissociative Identity Disorder) seem to confirm the idea of a fictional self. According to Dennett's interpretation it is sometimes rather disadvantageous for a cognitive system to attribute all experiences to one self alone. For the 'psychological survival' of some human beings the production of other selves might provide an advantage. The ontological status of these additional selves is the same as the one of the 'original self'—it is real as a fiction which is used by the system for specific tasks of self-representation.

In a similar vein as Dennett the prominent German neurobiologist Gerhard Roth argues. Roth considers the self as an illusion. Similar to Dennett, Roth refers to decentralized working processes of the human brain in order to support his illusion-thesis of the human self (Roth 2003, 394f.). In the brain there is nothing which corresponds—as a proper neural correlate—to the unity and centeredness of the psychological self as it is experienced in everyday life. For Dennett and Roth the assumption of personal identity (a robust concept of the self) has to be rejected because it leads to dualism which is not reconcilable with a scientific outlook of the world (Dennett 1991, 423). For Dennett there are only two possibilities to deal with the phenomenon of the human self: Either someone becomes a dualist and interprets the self as the manifestation of a spiritual, non-physical reality or someone is willing to pay the price for a less obscure and more scientific view of the matter and unmasks the self as fiction. Only if someone is willing to accept the self as a fiction of self-representing biological systems then the non-scientific assumption of a soul or "soul-pearls" (as Dennett puts it) becomes superfluous (Dennett 1991, 423).

2. ARGUMENTS AGAINST INCOMPATIBILISM

In their search for neural correlates of mental phenomena naturalists refer mainly to physical entities which can be subsumed under the ontological categories of events and processes: Dennett, e.g., refers to parallel

distributed processes in the brain (Dennett 1991, 187); Wolf Singer identifies "dynamically associated, synchronized cell assemblies" with the neural mechanism responsible for our impression of a unified conscious experience (Singer 2000, 134); for Edelman and Tononi "rapid re-entrant interactions among distributed neural populations" are the neural correlates for the unity of conscious experience (Edelman & Tononi 2000, 141); according to Flohr consciousness and self-consciousness are based upon neural processes, which are "mediated by the cortical NMDA synapse" (Flohr 2000, 255). Among those events and processes quoted above there are no proper physical correlates for our everyday assumptions concerning personal identity. Events, processes or physical states seem to be no good candidates for being proper physical correlates of assumptions which are connected with personal identity. We can ask, therefore, whether the situation changes if we take enduring entities as proper physical correlates for these mental phenomena.

A short analysis of contemporary naturalistic literature in philosophy of mind creates the impression that enduring entities are excluded a priori from the list of possible neural correlates of mental phenomena. It is the question, though, whether the denial of enduring entities as proper physical correlates of mental phenomena is essentially connected with a scientific account of human cognition or whether it is the consequence of a biased selection of physical correlates for mental phenomena.

In the search for proper physical correlates of mental phenomena we can distinguish entities which are *ontologically admitted* by a scientific theory T and entities, to whose existence the proponents of T are *ontologically committed*. The class of *ontologically admitted* entities include all entities which can be described in terms of T. Should it turn out that supposedly real entities are neither describable nor accountable in the framework of T, they must be regarded as unreal. In this sense, scientific theories admit or exclude specific entities. The class C_1 of entities which are ontologically *admitted* by a scientific theory T is not coextensive with the class C_2 of entities to whose existence the proponents of T are ontologically *committed*. By accepting T, I commit myself to the existence of determinate entities. Ontological commitments include only those entities whose existence is presupposed by T. An ontological commitment means that T is true only iff an entity x exists. The class C_1 of entities

which are admitted as legitimate objects of T is larger than the class C_2 of objects which falls under the category of ontological commitments of T (C_2 being a subclass of C_1).

We can classify the entities which are presupposed or admitted in scientific theories in the ontological categories of events, processes,[4] states, properties and that of enduring entities (like objects and things). In this case we get the following picture: There might exist a set of scientific theories which are *committed* only to the existence of events, states and processes: For quantum-theory, e.g., an ontological system consisting of events, processes, properties and states might be an adequate ontological framework. That does not mean, however, that the ontological *constraints* of quantum-theory imply that only events, properties, states and processes are admitted as legitimate objects of a scientific description. There exists also an adequate quantum mechanical description of enduring objects.

As a matter of fact most naturalists in the field of cognitive science and philosophy of mind do not refer to quantum mechanics but to neurobiology and related disciplines when it comes to the analysis of physical correlates of mental phenomena. Quantum mechanics is among the favourite scientific theories for dualists when it comes to defend the possibility of the mental interacting with the physical (Eccles 1994 and Hameroff 1994). They refer in their theories about the interaction of mental and physical phenomena to the 'non-deterministic' level of sub-atomic entities which is described by quantum mechanics. However, for these dualists there are, strictly speaking, no physical *correlata* for mental phenomena on the subatomic level but only physical *relata* with which the mental entities interact. The mental entities themselves do not have a physical correlate at all because they belong to a non-physical reality. Quantum physics is, though, not the kind of science which naturalists have in mind when they argue in favour of the naturalization of common sense phenomena like mental states, selves, and other folk psychological assumptions.

What are the ontological commitments of the scientific framework of neurobiology in which naturalists look for the physical correlates of mental phenomena? Unlike (perhaps) quantum-physics, neurobiology is not committed only to events, states, properties and processes. When we

---

[4]   For the distinction between events and processes see Steward 1997, 75ff.

classify the entities presupposed in neuroscientific theories according to the ontological categories mentioned above we find among them events, states, properties, and processes—like reactions to certain stimuli, brain states, and synchronized neural activities—but also enduring objects—like neurons and brain regions. As we have seen above, the class C_2 of entities which belongs to the ontological commitments of a theory T is a subclass of the class C_1 of the entities which are admitted by T as legitimate objects of scientific description. Therefore, when endurers belong to the subclass C_2 of a neurobiological theory T', it is trivially true that there are no ontological constraints connected with T' which exclude endurers from being also members of C_1—legitimate objects of scientific description for T'.

We can conclude that in the field of neurobiology and cognitive sciences there are no pressing grounds for a naturalist to consider only events and processes as physical correlates of mental phenomena. The ontological commitments of neurobiological theories comprise members of all ontological categories—events, processes, states, and enduring objects; and it is clear that there is no ontological constraint which would exclude enduring objects as legitimate objects of a neurobiological description. Why, then, are naturalists like Dennett et al. not willing to accept endurers as proper correlates of mental phenomena like the human self? One possible reason for this refusal might be that enduring neurobiological entities like neurons, brain regions are not good candidates for being proper correlates of the self because this would amount to assume a kind of *homunculus*. In this case, the proper correlate of the self would be some kind of "pontifical neuron" which controls all our brain activities (Dennett 1991, 413). As a matter of fact, the homunculus-view is presented by Dennett as the only candidate for a proper physical correlate of the human self. From the scientific implausibility of the homunculus-view he deduces that there is no proper physical correlate of the human self.

This conclusion, though, is premature. The class of endurant entities which are candidates for proper physical correlates of the self does not comprise only body parts like neurons or brain regions, but also functions. According to Barry Smith, functions fall under the ontological category of enduring entities: The function of our heart, for example, "begins to exist with the beginning to exist of your heart, and continues to exist, self-

identically, until (roughly) your heart ceases." (Smith 2004) In his formal ontology Smith follows Aristotle, who subsumes organisational principles under the category of substance (*ousia/forma substantialis*). Therefore, we have to extend our search for proper physical correlates of the self to the class of bodily functions.

It is clear that no single organic function can fulfil the requirements needed by a proper correlate of the human self. There is, however, a functional principle of the human organism, which can even be found in Dennett's naturalistic conception of consciousness and the self. In his *Consciousness Explained* Dennett assumes a "biological self" (Dennett 1991, 414). The biological self is given to all organisms that are able to distinguish (implicitly) between themselves and their environment (a capacity which Dennett ascribes also to protozoa). For Dennett, the biological self is "wired" in the biological structure of organisms (Dennett 1991, 427). The decisive point is that Dennett would hardly admit that the proper correlate for the conscious self of adult human beings is the biological self. He stresses the dichotomy between the biological and the conceptual self. Whilst the biological self is the organisational principle of a living being, the psychological self of adult human beings is only a conceptual construct. Contrary to other organisms the limits of a human self are not defined through the biological structure of the human organism. In contrast to the biological self the psychological self is not bound to its body, rather it has a life of its own. Under normal circumstances the slogan 'one self per body' is correct but in contrast to the biological self a clear correlation between psychological self and body is not possible. There are, however, scientific data which can be interpreted differently from Dennett's view:

(1) Dennett's proposal to consider the biological and the psychological self as inhomogeneous does not seem to be backed up by results from developmental psychology. George Butterworth explored extensively the origins of self-perception in infancy. In agreement with Dennett, Butterworth shows that a biological self (understood as an implicit knowledge about the limits and functions of one's own body) can already be noticed in an unborn fetus. Butterworth emphasizes, however, that there are no reasons to distinguish sharply between the biological self as basis of all proprioceptive activities and goal oriented behaviour on the one hand

and higher forms of self-conception on the other hand. These different kinds of a self are not incompatible with each other. Rather higher forms of self-conception presuppose the more basic ones. He talks of a continuum beginning with primitive ways of bodily self-perception and terminating with a mature concept of an autobiographical self in adult human beings:

> The point is that movement synergies reveal properties of the material self as an organized totality; species typical developmental processes will determine the extent to which such aspects of the categorical self become elaborated within higher order cognitive processes. (Butterworth 1992, 108)

According to this view the biological self is nothing else but an early form or the biological grounding of the psychological self.

(2) Further evidence, which obviously contradicts the incompatibilist view of most contemporary naturalists, can be found in the work of the neurobiologist Antonio R. Damasio. Damasio's findings can be used as an argument against the thesis of Dennett, Flohr, Singer et al. that physical correlates of higher forms of (self-)consciousness are to be found exclusively in the cerebral cortex. He shows that (self-)consciousness essentially depends on structures which belong to older phylogenetic areas of the brain which are closely interconnected with biological functions. Damage of parts of the diencephalon, the brainstem, or the upper part of the formatio reticularis leads to various forms of loss of consciousness. These structures are responsible for the regulation of basic living functions of the organism—the so-called *inner milieu*. The dividing line between the parts of the formatio reticularis whose damage leads to a change or loss of consciousness and those parts whose damage does not entail such consequences is quite clear (Damasio 1999, 236ff.). From the fact that these brain areas are essentially involved in control and representation of bodily processes, Damasio draws the conclusion that there is a direct connection between subjective experience, neuronal representation and the control of bodily processes. According to him, core-consciousness is immediately connected with permanent representations of fundamental organic functions. It is the so-called "proto-self" that makes this constant representation possible. Since these basic regulatory mechanisms are relatively stable, they provide an optimal foundation for referring to an identical subject, as it is presupposed in self-consciousness. A central

condition for the development of human subjectivity and self-consciousness is thus the representation of a dynamic equilibrium (homeostasis) of the various organic states through the proto-self. Self-consciousness arises, if an object, the organism, and the relation among the two are represented. The neurobiological basis of this proto-self is the representation of the causal relation between (interior and exterior) objects and the organism (emotions).

Damasio explicitly turns against a relativization of the self in terms of a mere fiction. His understanding of the self is not so much the consequence of a different understanding of mental phenomena. Rather Damasio (and Butterworth in his research) interpret bodily phenomena—as correlates of mental phenomena—in a different way from Dennett's, Roth's, Singer's and Flohr's account. The reality underlying the self is not limited to specific cortical processes. This does not mean that brain processes or events do not play a decisive role in the constitution of consciousness. They are crucial for the representation and cognitive processing of those fundamental bodily functions which keep an organism alive. The organizational and functional structure of the human organism and its multilevel mapping in the brain, though, are fundamental for the formation of a self (Damasio 1999, 144). Even if we change permanently throughout life, the structure and functional principle of our organism remain largely unchanged. Bodily processes are grounded in a unifying principle, which persists soundly from the beginning to the end of our life. Self-representation generates the impression of identity and unchangeability of a stable self because this invariant organizational principle of our organism is constantly represented as well (Damasio 1999, 141). This organizational principle is no fiction but a reality of our organism that controls fundamental bodily functions. According to Damasio, without this principle neither consciousness nor self-consciousness could arise because there would be nothing that maps what remains essentially the same throughout time.

> Moment by moment, the brain has available a dynamic representation of an entity with a limited range of possible states—the body. (Damasio 1999, 142)

Thus, there would be no basis on which our capacity to refer to ourselves could be grounded. This organizational principle that represents the constitution plan for our bodily structure is on a fundamental level the proper physical correlate of what we call personal identity. By pointing at functions and conservational processes in an organism which remain largely the same throughout a life span, the physical-bodily realm does not appear completely incommensurable with those mental phenomena which are not perceivable as events. Enduring entities like bodily systems, basic organic functions and organizational structures of the entire organism seem to be better candidates for proper physical correlates of mental phenomena as they are brain events or distributed neural processes.

3. CONCLUSION

In the understanding of many contemporary naturalists physical phenomena are conceived as being fundamentally different from mental phenomena: While we talk about processes and events in regard of the body (brain), people refer to identity, self and subjectivity in the mental realm. A disembodied conception of the mind stands vis-à-vis a pure event- and process-like conception of the body. At this point the reason for the naturalistic account of the self as a mere fiction becomes obvious. It consists in an event- and process-like conception of the body which leads to a dichotomy between mind and body. Some mental events, such as emotions, upshots of thoughts etc., may be reducible to physical events; but there are a whole set of mental phenomena that cannot be reconstructed as physical events, like subjectivity, mental dispositions or our notion of an "I". Once the route of distinguishing between bodily and mental phenomena, as sketched above, is taken, it is hard not to arrive at a result like Dennett's laid down alternatives.

Dennett's line of argument is a paradigm example of how the naturalistic preference for a restricted ontology in which there are only events, processes, states, and properties, but no enduring entities, shapes the relationship between the scientific approach to cognition and our folk psychological assumptions on personal identity and other mental phenomena. It can be shown that the naturalistic view of the self as an explanatory fiction is grounded in the inability to reconcile our folk

psychological intuition of personal identity with an event-ontological understanding of the human body.

The work of Damasio and other experts in the field of cognitive science indicates, however, that the physical correlates of states of consciousness are not assumed to be single neuronal events or complex neuronal activity patterns. Increasingly scientists become aware of the relevance of basic functional principles of the living organism and their multilevel neurobiological mappings for the generation of consciousness (see for instance De Preester et al. 2005).

In the ongoing discussions of cognitive science more and more complex organic unities and their interconnected functions, which can be subsumed under the ontological category of endurers, serve as proper correlates of states of consciousness. This implies that the correlated mental phenomena do not appear to be as strange and different as they are when physical correlates of the mental were sought exclusively in processes or events in specific regions of the human brain. Fundamental bodily functions together with their multilevel mapping in the human brain are good candidates for being proper correlates of mental phenomena like self-consciousness, subjectivity and our intuition of personal identity: Bodily functions remain the same even when they are not actualized. They are dispositional in character and thus, not adequately understandable within a framework of mere event-ontology. A comprehensive understanding of the totality of scientific data requires a richer ontology which comprises not only states, processes and events but also enduring continuants (substances). If everything is like a process or event, we have not only an inadequate understanding of our common sense intuitions, but we are also not able to understand the whole range of scientific data in a comprehensive way. An adequate understanding of the structure of the human body, the human mind and its working together needs both—continuants and events.

4. REFERENCES

Butterworth, G. 1992: "Origins of self-perception in infancy." *Psychological Inquiry* 3, 103-111.

Chalmers, D. 1995: "Facing up to the Problem of Consciousness." *Journal of Consciousness Studies* 2, 200-219.

Churchland, P. M. 1979: *Scientific Realism and the Plasticity of Mind.* Cambridge, MA: Cambridge University Press.

Churchland, P. M. 1990: "Eliminative Materialism and the Propositional Attitudes." In: Lycon, W.G. (ed.) *Mind and Cognition. A Reader.* Oxford: Blackwell, 206-223.

Damasio, A. 1999: *The Feeling of What Happens. Body and Emotion in the Making of Consciousness.* New York: Harcourt Brace.

Dennett, D. 1991: *Consciousness Explained.* Boston: Little, Brown and Company.

De Presster, H. and Knockaert, V. (eds.) 2005 *Body Image and Body Schema. Interdisciplinary perspectives on the body.* Amsterdam: John Benjamins Publishing Company.

Eccles J C. 1994: *How the Self Controls its Brain.* Berlin, New York: Springer.

Edelman, G. M. and Tononi, G. 2000: "Reentry and the Dynamic Core: Neural Correlates of Conscious Experiences." In: Metzinger, T. (ed.) 2000, 139-151.

Flanagan, O. 1992: *Consciousness Reconsidered.* Cambridge, MA: MIT Press.

Flohr, H. 2000 "NMDA Receptor—Mediated Computational Processes and Phenomenal Consciousness." In: Metzinger, T. (ed.) 2000, 245-258.

Goldman, A. I. 1970: *A Theory of Human Action.* Englewood Cliffs: Prentice-Hall.

Hameroff, S. 1994: "Quantum coherence in microtubules: a neural basis for emergent consciousness?" *Journal of Consciousness Studies* 1, 91-118.

Jansen, L. 2007: "On Ascribing Dispositions." In: Kistler, M. and Gnassounou, B. (eds.) *Dispositions and Causal Powers.* Aldershot: Ashgate, 161-177.

Lewis, D. 1966: "An Argument for the Identity Theory." *Journal of Philosophy* 63, 17-25.

Metzinger, T. 1993: *Subjekt und Selbstmodell. Die Perspektivität phänomenalen Bewußtseins vor dem Hintergrund einer naturalistischen Theorie mentaler Repräsentation.* Paderborn: Schöningh.

Metzinger, T. (ed.) 2000: *Neural Correlates of Consciousness. Empirical and Conceptual Questions.* Cambridge, MA: MIT Press.

Mumford, S. 1998: *Dispositions*. Oxford: Clarendon Press.

Parfit, D. 1984 *Reasons and persons.* Oxford: Oxford University Press.

Quine, W. V. O. 1976: "On Multiplying Entities." In: Quine, W. V. O. *The Ways of Paradox and other Essays.* Cambridge, MA: Harvard University Press.

Roth, G. 2003: *Fühlen, Denken, Handeln. Wie das Gehirn sein Verhalten steuert.* Frankfurt am Main: Suhrkamp.

Singer, W. 2000: "Phenomenal Awareness and Consciousness from a Neurobiological Perspective." In: Metzinger, T. (ed.) 2000, 121-137.

Smith, B., Munn, K. and Papakin, I. 2004: "Bodily Systems and the Spatial-Functional Structure of the Human Body." In: Pisanelli, D. M. (ed.) *Ontologies in Medicine: Proceedings of the Workshop on Medical Ontologies, Rome, October 2003*. Amsterdam: IOS Press, 39-63.

Steward, H. 1997: *The Ontology of Mind.* Oxford: Clarendon Press.

Von Eckardt, B. 1993: *What is Cognitive Science*? Cambridge, MA: MIT Press.

The Unmysteriousness of Consciousness: A Case Study in Naturalistic Philosophy

Johannes L. Brandl
University of Salzburg

1. INTRODUCTION

Naturalism in philosophy of mind is commonly associated with a materialistic or physicalistic conception of reality. This association is often made because prominent physicalists, like David Armstrong and Willard Van Orman Quine, assume a naturalistic perspective in defending their metaphysical views. Both Armstrong and Quine acknowledge, however, that these doctrines are independent of each other at least in the following sense: if physicalism should turn out to be an untenable position, this would not force one to give up one's naturalistic outlook in philosophy in general (Armstrong 1980, 156ff; Quine 1995, 257).

In this paper, I want to make an even stronger claim. I think that there is a tension between the spirit of a naturalistic philosophy and the metaphysical commitments made by a physicalist. Associating these views means ignoring this tension and can mislead one into thinking that a monistic view is more congenial to naturalism than a dualistic metaphysics. In fact, however, it is hard to find good reasons for such an asymmetry. If this is so, as I shall argue, naturalists should adopt a stance that is metaphysically neutral in this respect.

To make this point, I will use the problem of consciousness as a case study. David Chalmers has urged that in explaining consciousness we must distinguish between comparatively 'easy problems' connected with a functionalist conception of the mind, and a 'hard problem' that arises in explaining phenomenal consciousness (Chalmers 1995, 1996). This distinction has provided new support for the claim that consciousness is in

some sense a deeply mysterious phenomenon (McGinn 1989, 1999).[1] According to this view, consciousness is a mystery that is in no way reduced by the advances of psychology and the neurosciences. Du Bois Reymond may have been right, when he proclaimed his famous *Ignorabimus* about consciousness in a lecture on the limits of natural science in 1872 (reprinted in Du Bois Reymond, 1974).

Naturalism, as I understand it, is strictly opposed to such skepticism because naturalism does not allow philosophy to pass judgment on what the sciences can or cannot achieve. Declaring consciousness a mystery is just as mistaken as believing in unlimited scientific progress. Naturalists can accept neither of these claims. Their view has to be that the sciences determine their own path without being constrained by *a priori* philosophical reasoning.[2]

As one can see, the topic of this paper is huge and has many ramifications. In dealing with it here, I will have to set aside many issues for further treatment. The brevity of my remarks, I hope, will help in grasping the larger picture that I try to paint here. The paper is organized as follows. In section 1, I first consider briefly the historical roots on which a mysterian conception of consciousness relies, then I set out a contemporary argument that supports this view, and I consider several ways how one can resist its conclusion. The naturalistic reaction to this argument, I suggest, should be to question the alleged metaphysical implications contained in scientific theories of the mind. In section 2, I will develop this idea further by tracing it back to the anti-metaphysical stance of logical empiricism. At this point, the question will arise how this stance can be squared with the fact that the members of the Vienna Circle were drawn towards a

---

[1] Chalmers himself draws a different conclusion. He thinks that a scientific explanation of phenomenal consciousness may well be possible, but it will require theories of a quite different kind than those currently on offer. The task of philosophy is to pave our way to a kind of 'conceptual breakthrough' that would enable us to take experiences as basic features of reality alongside mass and gravity.

[2] Although Du Bois Reymond was a physiologist, his argument that consciousness will remain forever a mystery was not an empirical prediction, but a philosophical point based on *a priori* reflections "on the nature of things" (Du Bois Reymond, 1974, 65). For a detailed study of his lecture and the impact it had in the 19th century, see Vidoni 1991.

physicalistic ontology. This move finds its completion in Quine's philosophy, as I will explain in section 3. Whereas Quine seems inclined to sacrifice the metaphysical neutrality of his predecessors, I will suggest that this neutrality should be retained. In section 4, I sketch a modest version of naturalism that is built on this idea. In section 5, I show how this modest approach helps to overcome the sense of mystery surrounding the problem of consciousness; and in section 6, I try to meet two objections that this modest naturalistic approach has to face.

2. THE MYSTERY OF CONSCIOUSNESS

The fact that we experience the world in a subjective manner can seem very puzzling. Philosophy articulates this puzzlement in a systematic form and tries to come to terms with it. Within this tradition, from Plato onward, it was widely assumed that the source of this puzzlement is a metaphysical one. The problem is to understand the relation between mind and body: between the subjective experiences that fill our minds when we are conscious and the physical powers that move our bodies. But the mind-body problem is not the only reason why the occurrence of subjective experiences seems such a curious fact.

Experiences are also peculiar because we cannot perceive them in the way in which we can perceive bodily things. How, then, can we use the same methods in psychology that we use in the natural sciences? This became an urgent question in the 19th century when psychology developed into an experimental science. Philosophers like Brentano and Husserl reacted to this new situation by distinguishing two kinds of psychological research: there is psychology as a natural science, and there is psychology as a philosophical or phenomenological discipline (Brentano 1982, Husserl 1950). The former requires the study of physiological processes and behavioral reactions causally connected with conscious experiences; the latter requires inner perception or a phenomenological epoché.

The idea of splitting psychology in two halves opened up a gap that troubles philosophy until today. The first problem here is to understand the nature of this gap. Is it just a methodological distinction between disciplines that use different methods in studying the same subject matter? Or is there also a deeper ontological difference involved here? The

distinction between a functionalist and a phenomenological conception of mind has made this problem virulent again (Chalmers 1996). Most agree that this is a valuable *conceptual* distinction that reflects the important difference between a first-personal and a third-personal access to our mental states. Whether there is also a more fundamental metaphysical difference to be drawn here between functional and phenomenal *states* or *properties*, as Levine and others have suggested, remains a matter of dispute (Levine 1983).

This was a short summary of what lies behind the skeptical views about explaining consciousness that I want to consider now. In setting out the problem, I will adopt the jargon of speaking about qualia and qualia differences. However, I do not think that much depends on this terminology and consider it merely a convenient device for shortening the argument. As I use the term here, qualia are properties of experiences, and I assume that experiences must have qualia if they are instances of phenomenal consciousness. Sensory experiences are the prime examples of mental events instantiating qualia. For instance, when one sees some ripe strawberries, smells them and then tastes them, one perceives the same object in three different ways. These perceptual experiences differ qualitatively from each other, which means that they carry different qualia.

The qualia problem can be phrased in two different ways: (1) How, if at all, can the fact that experiences have different qualia be explained within a physicalistic conception of the world? (2) And how, if at all, can this fact be explained within empirical science? The first question is a metaphysical, the second a methodological one. Thus we see that both of the traditional concerns about experience continue to play a role here. How they interact with each other and thereby get merged into a single problem, is a complex question that I cannot pursue here any further (van Gulick 1996).

Let me turn directly to the argument in support of the view that consciousness is a mysterious and inexplicable property of mental states.

The argument can be summarized in five steps:

(1) Qualia are real properties of conscious experiences.
(2) The existence of qualia is compatible with a physicalistic conception of reality only if qualia differences can be fully explained in terms of physical (primarily neurophysiological) differences.
(3) Qualia differences cannot be fully explained in terms of physical differences.
(4) Science is committed to a physicalist conception of reality.
(5) If science cannot fully explain the existence of qualia, their existence is mysterious.

This is a sophisticated argument that combines both methodological and metaphysical considerations. The two aspects are linked in premise (2), which introduces the requirement of a 'full' explanation as being an explanation that satisfies a physicalist. Step 4 makes this notion of explanation the goal of science, from which it follows – together with premise 3 – that qualia differences cannot be explained within science. With premise 5, and *modus ponens*, the skeptic can derive his conclusion that the existence of qualia is mysterious, and with premise 1 he can finally conclude that conscious experience is mysterious too.

Each step in this complex argument provides an opportunity for resisting the skeptical conclusion. Which of its premises one rejects, depends on what kind of solution to the mind-body problem one prefers. Let me briefly review the three options that recently have been debated most widely.[3]

One option is to reject the very first premise of the argument by denying that qualia are real properties of experiences. This would require one to argue that there is a fundamental problem in the conceptual framework employed in distinguishing between a functional and a phenomenal conception of the mind. The claim would have to be that we do not need this distinction or any similar distinction between different kinds of

---

[3] For a more comprehensive survey of some recent attempts at providing reductive and non-reductive solutions for the mind-body problem, see Van Gulick 2001.

consciousness, if we want to explain our cognitive engagement with the world. This line is taken, for instance, in (Rey 1983) and in (Wilkes 1984).

A second option is to argue that qualia may find a place within a physicalistic worldview even though they are not reducible to physical properties. They may not even be strongly supervenient on physical properties and hence it may be impossible to trace qualitative differences to physical differences. Non-reductive physicalists who advocate this view have to come up with some other explanation of how qualia fit into the physicalist's conception of the world. Alternative explanations that should also satisfy a physicalist have been proposed, for instance, in (Tetens 1996) and (Van Gulick 2002).

A third option is to claim that qualia differences can be traced back to physical differences and fully explained in terms of them. Only a lack of empirical knowledge has prevented us so far from seeing the right connections here. As more of this knowledge becomes available, we will gradually be able to work out a theory of psychophysical supervenience, or psychophysical reduction. A project of this kind that pays close attention to the empirical progress made by the neurosciences, has been suggested for instance in (Bickle 1998).

All these proposals have one thing in common, which is also their common weakness: they are all extremely demanding replies to the skeptical argument. None of these envisaged projects has been carried out so far, and it is not clear which of them is the most promising one. This is, of course, water on the mills of the skeptic. He will take this fact to be further evidence for his claim that the problem of explaining phenomenal consciousness is simply too hard for the human mind to solve.

But we are not quite finished with listing all the available options. A further option would be to deny premise (4).[4] That this option is often overlooked is not surprising, since physicalists share this premise with their opponents. They also assume that science is committed to such a metaphysical position, they only draw different conclusions from this premise. This is an unfortunate agreement, I think. It deprives us of an easy

---

[4]   Denying premise (5) would be another option that I set aside here. Those who believe in transcendental explanations may want to take this route. In the final section of this paper, I will briefly indicate why this possibility must be dismissed from a naturalistic point of view.

way to bring the skeptical argument to a halt, namely by invoking the difference between advocating naturalism and advocating physicalism. Before we can count on this move, however, we must see what this difference comes to.

3. NATURALISM, PHYSICALISM, AND LOGICAL EMPIRICISM

Naturalism in philosophy[5] is a broad movement with many different currents in ethics, epistemology and metaphysics. This makes it difficult to say in general what this view includes and what it denies. Definitions of naturalism therefore tend to be rather uninformative. An example is Roy Wood Sellars' statement that naturalism is the

> recognition of the impressive implications of the physical and biological sciences. (Sellars, 1922, i)

This is an unsatisfactory explanation for two reasons: first, it describes naturalism as an attitude without saying on which principles this attitude is based; and secondly, it describes an attitude that almost everyone shares. There is hardly a philosopher since the enlightenment who would not recognize the important contributions of the physical and biological sciences. Naturalism in this sense ceases to be a controversial and interesting philosophical position.[6]

   The situation changes, however, when we consider the historical context of Sellars' statement: the rise of logical empiricism, which soon became one of the most controversial movements in 20th century philosophy. Although the term 'naturalism' was not widely used by the members of this school, they clearly embraced the attitude that Sellars advocates. There are two principles of logical empiricism that deserve special attention here: the rejection of metaphysics and the proposal of a common language for all sciences.

---

[5] Throughout this paper I use the term "naturalism" as short for "philosophical naturalism", i.e. a movement within philosophy, not the arts.

[6] This is a problem not only for Sellars' characterization, but affects many of the slogans which are often taken to be expressions of a naturalistic point of view. See Keil/Schnädelbach 2000.

When Carnap and other members of the Vienna Circle first put forward these claims, they chose to give them a most dramatic formulation. They declared all metaphysical statements to be "senseless pseudo-statements" (Carnap 1931a), and they claimed that all scientific statements could be "translated" into a physicalistic language (Carnap 1931b). On closer examination, however, the real content of these provocative claims turns out to be far less radical. We might capture it in the following way:

(P1)   Metaphysical statements that transcend science have no explanatory value.

(P2)   A unifying bond for all scientific theories will be the use of a quantitative language in describing their evidential base.

I will restrict my discussion of these principles here to the role they play in dealing with the mind-body problem. This problem was of major concern to many logical empiricists, and they approached it on the basis of the two principles stated above.

The first thing to note about their approach is that it was not original (Heidelberger 2003). It followed a popular trend in *Naturphilosophie* in the 19th century, one of whose main representatives was the physiologist Gustav Theodor Fechner. In his widely acclaimed book *Elements of Psychophysics* (1860) Fechner advocated a position called "psychophysical parallelism". This is a view that admits of different interpretations. One reading of it leads to Fechner's panpsychism, another interpretation leads to a mind-brain identity thesis. Underlying these metaphysical claims, and supporting them, is an empirical thesis that Heidelberger states as follows:

> The primary form of psychophysical parallelism is an empirical postulate – a methodological rule for researching the mind-body relation, claiming that there is a consistent correlation between mental and physical phenomena. [...] This type of psychophysical parallelism refrains from all causal interpretation of the mind-body relation. Fechner said that it is neutral regarding every "metaphysical closure" compatible with it. (Heidelberger 2003, 237)

As Heidelberger notes, there is a remarkable similarity here to what William James says 30 years later:

William James [...] confined himself – as he said – to 'empirical parallelism' [...].
'By keeping to it', he wrote in *Principles of Psychology*, 'our psychology will
remain positivistic and non-metaphysical; and although this is certainly only a
provisional halting-place, and things must some day be more thoroughly thought
out, we shall abide there in this book (James 1891, 182). (Ibid, 238)

Both Fechner and James believed that the empirical sciences could provide
at least a partial solution to the mind-body problem. They can do this by
tracing qualia-differences back to physiological differences, even if they do
not thereby provide a 'full' explanation of this connection. A 'full'
explanation would be – in accordance with premise 2 of the skeptical
argument considered in the previous section – an explanation that satisfies
the demands of a physicalist.

Logical empiricism gave this idea a different twist by taking a critical
stance with respect to metaphysical explanations in general. From their
point of view, the only interesting aspect of the parallelism thesis is that
mental events co-occur with physiological events in the brain. No further
interpretation of this parallelism is needed, as Fechner thought; and no
further insights are to be expected, when "things are thoroughly thought
out". This is the point of the anti-metaphysical principle (P1).

A letter to Cassirer that Schlick wrote in 1927, quoted by Heidelberger,
shows where their opinions deviated here:

The psychophysical parallelism in which I firmly believe is not a parallelism of
two 'sides' or indeed 'ways of appearing' of what is real, rather, it is a harmless
parallelism of two differently generated concepts. Many oral discussions on this
point have convinced me (and others) that in this way we can really get rid of the
psychophysical problem once and for all. (Ibid. 250)

It may seem, however, that James was right after all and that the 'harmless
parallelism' that Schlick adopts here was just a 'provisional halting-place'.
A few years after this letter was written, physicalism became the official
doctrine of the Vienna Circle. Did the logical empiricists finally realize
that the psychophysical problem required a more profound solution? This
depends on what their move to 'physicalism' involved.

The usual way to understand this move is to classify it as a 'linguistic'
or 'semantic physicalism' whose goal is to eliminate problematic non-
physical entities from psychological theories by expressing them (or

'translating' them) into the language used in the physiological and behavioral sciences. But this is a highly questionable reading of the view embraced by logical empiricism at this time. A very different reading has been suggested by Stubenberg who interprets this view as a form of "non-materialistic physicalism" (Stubenberg 1997, 144). This sounds more paradoxical than it actually is. The crucial point, I think, is that the move to physicalism was – at least initially – a response to a methodological question: what is the evidential basis of scientific theories? Following Ernst Mach, Schlick and Carnap had felt comfortable with assuming that the evidence of scientific theories consists in experiences that directly verify statements like 'there is a red spot', and 'it is cold now'. This assumption was challenged by Otto Neurath who claimed that these qualitative statements can, and should, be replaced by reports using only quantitative terms like 'light of length L is reflected there' and 'the kinetic energy reaches level L now'. Neurath's point was not a metaphysical, but a pragmatic one: statements of the first kind carry no additional *evidential weight*, whereas statements of the latter kind have the advantage of introducing quantitative methods into all sciences. They can therefore serve as a unifying bond between the natural and the social sciences, including psychology and the humanities. This is the point of the unity principle (P2).

Thus conceived, the move to 'physicalism' within logical empiricism has not much to do with the mind-body problem. It was a move within epistemology, not within metaphysics. This may also explain another puzzling fact that Jaegwon Kim noticed recently. He pointed out that even in classical texts from this period, like Hempel's paper "The Logical Analysis of Psychology" (1935), one does not find a clear commitment to a physicalist position:

> This means that Hempel's translatability thesis – the claim that all psychological statements are translatable into physical statements – is fully consistent with the Spinozistic or Leibnizian dualism. [...] The conclusion seems inescapable that the notion of translation used by Hempel [...] cannot serve as a basis for formulating a robust and significant form of physicalism. (Kim 2003, 268)

Logical empiricists could therefore still reject metaphysical solutions of the mind-body problem as nonsense. This caused no tension for them, because

their physicalism was not a metaphysical position in the first place. It was just a methodological rule about how to formulate the evidential base on which scientific theories rest. But the tension did arise eventually. There is no doubt that logical empiricism was driven towards a physicalistic position worth this name, even if their metaphysical commitments remained unclear and changed between a functionalist theory and a psycho-neural identity theory. The anti-metaphysical stance that characterized their view initially thereby went over board.

In Kim's view, this was a salutary move in the right direction (ibid, 275 and 277). However, when one takes into account the qualia problem, this judgment may have to be reconsidered. Neither functionalist theories nor identity theories are very successful strategies in answering the skeptical argument explained earlier. Perhaps logical empiricism took a wrong turn here, when it eventually followed Fechner in adopting a more robust physicalistic theory of the mind. In the next section, I shall provide further evidence for this hypothesis by turning to Quine's naturalism.

4. QUINE'S NATURALISM

The prominent role that Quine has played in promoting naturalism in contemporary philosophy cannot be missed. He is widely recognized as the main representative and the spearhead of a naturalized epistemology and a naturalized metaphysics (Kitcher 1992, Craig/Moreland 2000, DeCaro/ Macarthur 2004). In the same way in which logical empiricism built on the ideas of the 19th century, Quine's naturalism was derived from his predecessors. Although this fact has been widely acknowledged in the literature, it has not been fully appreciated how selective Quine was in continuing the tradition of logical empiricism. For Quine, naturalism consists in

> the recognition that it is within science itself, and not in some prior philosophy,
> that reality is to be identified and described. (Quine 1981, 21)

This echoes not only Roy Wood Sellars' statement quoted earlier, but also the denial that metaphysics can (or must) go beyond science. If one recognizes the importance of the natural sciences, Quine says here, one

must also acknowledge that science takes care of the most general questions we can ask about reality. That is to say, it also takes care of our metaphysical problems, like the mind-body problem. Philosophical views on these matters are legitimate only to the extent that they stay within the boundaries set by the empirical sciences.

How tight are these boundaries, and how are these boundaries fixed? Quine's answer to this question is complex, and it is beyond the scope of this paper to deal with it in detail. The following remarks can only provide a rough sketch of Quine's position.[7]

The claim that experience limits metaphysics is an old claim of empiricism. However, as Quine argues, both the classical version that we find in British empiricism and the version we find in logical empiricism have to be updated in order to bring this claim in line with contemporary science. Two assumptions have become untenable for Quine: the assumption that sensory evidence consists in impressions that give rise to ideas; and the assumption that sensory evidence is distributed and associated with single sentences or even single ideas. Quine's project is to develop a form of empiricism that is purified of mental concepts, like the concept of an idea, and holistic in taking science as a whole to face the tribunal of sensory evidence (Quine 1981b, 67ff).

Where does naturalism come into play here? It enters the scene in the way in which Quine tries to get rid of the mentalistic vocabulary:

> [...] my stance is naturalistic. By sensory evidence I mean stimulation of sensory receptors. I accept our prevailing physical theory and therewith the physiology of my receptors, and then proceed to speculate on how this sensory input supports the very physical theory that I am accepting. (Quine 1981, 24)

But, does Quine really speak as a naturalist here, as he claims, or as a physicalist? This depends on what one means by 'physicalism'. If one uses this term in the way in which logical empiricists used it when they requested that the evidence supporting scientific theories should be described in a quantitative language, then the question makes no difference. Quine's physiological account of sensory experience is clearly in line with this request and therefore both naturalistic and physicalistic in

---

7   A fuller treatment of Quine's position can be found, for instance, in (Hylton, 1994).

this wider sense. But why isn't it also physicalistic in the stronger sense in which physicalism is a doctrine about the mind-body relation? And if it is, why does Quine not say so instead of describing his position as merely 'naturalistic'?

The answer, I suspect, has to do with Quine's view that his claim about the nature of sensory evidence can be derived from empirical science and needs no extra metaphysical foundation. It is supposed to follow from the "prevailing physiological theory of our sense receptors". But is this appeal justified? Even if neurophysiology has much to say about the processes by which we pick up information via our senses, this does not show that there is nothing more to sensory experience than neurophysiological processing. This is something that Quine tacitly assumes here, and in assuming it he smuggles a metaphysical premise into his empiricism.

I should make clear that I am *not* criticizing here Quine's commitment to a physicalist theory of the mind. Whether physicalism provides a correct account of mental states is not the issue here. I am concerned with Quine's claim that he is speaking as a *naturalist*, not as a physicalist, when he identifies sensory experiences with stimulations of sensory receptors. He thereby suggests that his naturalism is something more fundamental and independent of his physicalistic commitments. In fact, however, his naturalism seems to have absorbed his physicalist commitment.

We can now also understand why Quine requests that the language of science should be purified from any mentalistic idiom. In making this request, Quine is not merely modernizing empiricism; he is confronting a Cartesian epistemology and metaphysics. This becomes most vivid when Quine replaces Descartes' description of himself as a reflecting *res cogitans* with his own physicalistic self-description:

> I am a physical object sitting in a physical world. Some of the forces of this physical world impinge on my surface. Light rays strike my retinas; molecules bombard my eardrums and fingertips. I strike back, emanating concentric airwaves. These waves take the form of a torrent of discourse [...]. (Quine 1966, 228)

This is a *philosophical* statement. There is no need in science to adopt such a point of view, as one can see from the fact that mentalistic terms are used all over in psychology and the social sciences. Quine criticizes this practice

and offers a rational reconstruction of science that shows how mentalistic expressions may be eliminated. But such criticism needs an extra-scientific basis. It has to be grounded in a metaphysical doctrine that is not contained in science itself.

Quine's naturalism, I said earlier, absorbed his physicalism and thereby lost its metaphysical innocence. It is not content with the demand, originally introduced by Neurath, that the evidential base of science should be described in a quantitative (i.e. 'physcialist') language that may serve as unifying bond for science. It demands in addition that the evidence should be *exhaustively* described in such a language. No qualitative or intentional terms for marking qualitative differences among experiences are allowed. This stronger claim rests on a full-blown physicalism, not just a "non-materialistic" variant of it.

5. FROM PHYSICALISM TO MODEST NATURALISM

As the previous section has shown, the difference between naturalism and physicalism that Quine officially recognizes becomes very thin when one takes into account his anti-mentalistic stance. But the distinction is still worth drawing. In this section, I therefore suggest a change of course. I think we should leave behind the anti-mentalistic stance of Quine and instead return to the goal of metaphysical neutrality initially pursued by logical empiricism. This will make room for a modest form of naturalism and a broader notion of evidence that does not reduce to informational processes taking place at the sensory and neurophysiological level.

Modest naturalism, as I understand it, is a pluralistic doctrine according to which knowledge can arise from many different sources. There are the sources that can be explained in terms of chemical or biological processes, but there are other sources as well that are therefore no less 'natural'. The first task for a modest naturalist, therefore, is to introduce a broader notion of what it means to be 'natural' that is not tied to the perspective of the natural sciences. I suggest that we understand this notion in the following way:

> A property F is natural in a broad sense if a reasonable explanation can be given as to why objects exemplify F.

Reasonable explanations come in many different forms. There is a reasonable explanation why objects have a certain mass, why they are soluble, and why they function as kidneys or hearts. These explanations are given by physics, chemistry, and biology respectively. But there is also a reasonable explanation why certain objects are linguistic symbols, presidents, or pieces of art. These explanations are given in linguistics, in social studies, and in aesthetics. This distinguishes them from all non-natural properties, like being an angel or a work of witchcraft, which cannot be explained in this way. Even if one can tell stories about such things, they remain a mystery to us and therefore do not count as 'natural' even in a broad sense. Using this wider notion, we can now say more generally what a natural source of knowledge is:

> A source of knowledge is natural in a broad sense if a reasonable explanation can be given as to how someone can acquire knowledge from this source.

One possible explanation for acquiring knowledge starts at the sensory level and explains how we pick up information from stimulations on the surfaces of our body. But much of what we know cannot be explained in this way. When we read something in the newspaper, the source of our knowledge is much richer than just a pattern of sensory stimulations: it consists of the entire set of practices and institutions that is needed for a newspaper to function properly. There has to be a written language and a tradition of spreading reliable information via such media.

The basic principle of modest naturalism can now be stated as follows:[8]

(N)   All knowledge about the empirical world is derived from natural sources of knowledge, i.e. from sources whose functioning can be reasonably explained.

---

[8]   Different versions of a modest (or harmless) naturalism have been proposed by John McDowell (McDowell 1996, 2004), Jennifer Hornsby (Hornsby 1997), Robert Almeder (Almeder 1998) and Hans Fink (Fink 2006). A comparison of these views with the one developed here is beyond the scope of this paper.

Is there anything controversial about this principle? This is a tricky question that one can answer both with "Yes" and "No". Since the present view agrees with the basic idea of logical empiricism that metaphysical statements have no explanatory value when they transcend science, it has to be controversial. But, on the other hand, it seems to express a conceptual truth. It merely tells us that an epistemic state of belief can count as an instance of knowledge only if it is justified in terms of a source that can be reasonably explained. How could this be denied?

We must therefore conclude that conceptual truths can also have controversial consequences. I think this is as it should be. Conceptual analysis, after all, is not a tedious enterprise of merely making explicit what everyone knows and no one doubts. As a conceptual truth, (N) does not need support from an extra metaphysical premise. It therefore also preserves the goal of remaining metaphysically neutral. I now want to show how this feature of modest naturalism can help us to diffuse skepticism about explaining phenomenal consciousness.

6. DEMYSTIFYING CONSCIOUSNESS

In section 1, I distinguished between different forms of puzzlement to which consciousness can give rise. There are mysteries that originate from the traditional mind-body problem, and there are mysteries that result from the diverging methodologies of phenomenology and empirical psychology. A mixture of both sources provides the background for a radical skepticism that agrees with Du Bois Reymond's *Ignorabimus* claim. The argument leading to this conclusion was that there is no hope that the qualia of experience can be integrated into a physicalistic world-view, nor is there hope that they could be explained without such integration or that they could be successfully eliminated from a psychological description of mental reality. This leaves phenomenal consciousness as a 'riddle' that cannot be explained away.

But there is a loophole in the argument, as I pointed out. In fact there are *two* loopholes in the assumption that a 'full' explanation of consciousness has to be both physicalistic and scientific. The modest naturalism I am advocating here accepts the latter part of this assumption: an explanation of consciousness that does not meet the standards of a

scientific explanation could not count as a 'full' explanation. This does not force us to accept also the first part of this assumption. An explanation could count as a 'full' explanation even if it does not meet the requirements set up by a physicalist. Using the broader notion of what it means to be a natural phenomenon, a naturalist will claim no more – and no less – than the following:

(CN)   Consciousness is a natural property in the sense that there is a reasonable   explanation how living creatures come to have conscious experiences.

The term 'conscious experiences', as I use it here, covers all types of conscious mental states, be they sensory experiences resulting from perception and memory, experiences of desire or emotion, or experiences we have in problem solving or in forming abstract thought. A distinction between functionalist and phenomenal types of consciousness plays no role at this level, although it may be introduced later when different explanations are given for mental states instantiating these different kinds of consciousness. The distinction then presupposes that *both* types of consciousness can be explained. It therefore cannot be used for distinguishing between a form of consciousness that can, and another that cannot be reasonably explained.

Although (CN) is a very modest claim, it needs to be defended against the view that consciousness is mysterious and therefore not a natural phenomenon. There are two possible routes a modest naturalist can take here: he may choose a 'deflationary' defense or a 'dialectical' defense of his position.

Taking the deflationary route, he could point out that fragments of a theory of consciousness are already available. Neurophysiology has shown which parts of the brain are active when certain experiences occur; cognitive psychology has developed models that may explain how informational states become conscious when they are globally broadcast in the brain; and social studies have shown how social interaction and cultural activities influence the development of advanced forms of self-consciousness. These fragments give us already a partial understanding of

he emergence of consciousness in human and nonhuman animals. Future research will make this picture more and more complete.[9]

This line of defense shows how much a skeptic has to set aside when he declares consciousness to be an irresolvable mystery. But pointing this out will not be enough to win the battle. A skeptic can respond here that the partial understanding provided by our current scientific theories creates an illusion. It is the illusion that we only need better theories of the same kind to complete the job. But theories of the same kind will not be able to dissolve the mystery, the skeptic may insist, and therefore the glass will always remain half empty.

In order to overcome a skeptical challenge of this sort a dialectical response is needed, and I think that a modest naturalist is in the best position to offer such a response. He can point out, as we have seen, that his position is neutral and involves no commitment to a form of physicalism or dualism. Therefore, he need not solve the difficulties that physicalists and dualists confront in dealing with the mind-body problem. He can thus set aside the problem how qualia could be integrated into a physicalistic conception of the world, and how mental states can cause physical states, and vice versa, on a dualistic theory. The skeptic may be right that *these* are intractable problems. But the problem how consciousness arises in living creatures is not to be identified with any of these problems and therefore can be solved. In sticking to his metaphysically neutral position, the naturalist can thus turn the tables against the skeptic.

7. TWO OBJECTIONS

Can a naturalist get away with this reply to the skeptic? In concluding this paper I want to consider two objections that a modest naturalism has to face here. The first objection is that his position is *unstable*; the second that it is philosophically *self-destructive*.

---

[9]   This seems to be the view proposed in (Mills, 1996), and it may even be a view that McGinn finds attractive (see McGinn 1989 and 1999).

The objection that a modest naturalism, as I have described it here, is an unstable position has been voiced by Barry Stroud. He puts the objection in the form of a dilemma:

> There is pressure on the one hand to include more and more within our conception of 'nature': so it looses its definiteness and restrictiveness. Or, if the conception is kept fixed and restrictive, there is pressure on the other hand to distort or even to deny the very phenomena that a naturalistic study is supposed to explain. (Stroud 1996, 44)

Stroud thinks that some compromise has to be found here. A naturalist must accept some restrictions on what counts as natural and what not:

> Some determinate conception of what the natural world is like is needed to give substance to the claim that one's epistemology, or one's study of any other aspect of the world, is naturalistic. (Ibid, 45)

If anything whatsoever counts as 'natural', one has simply deprived naturalism of its content.

My response to this objection is that metaphysical neutrality does not mean 'anything goes'. It is quite plausible that *some* constraints will have to be made to sustain the claim that consciousness is a natural phenomenon. For instance, we have to assume that conscious experiences are datable occurrences in order to correlate them with neurophysiological activities in the brain. But this is a claim that does not go beyond a weak empirical parallelism, as explained in section 1. It still allows us to deny that angels or ghosts could also be conscious beings, something that we could not reasonably explain.

The metaphysical neutrality of naturalism that I am advocating here is a neutrality about how we interpret mental predicates like 'tasting ripe strawberries'. We can interpret them as denoting an irreducible mental property, like a property dualist does, or we can take them to denote some physical state – we do not know which – of our sensory system. But we need not commit ourselves one way or the other, since an *explanation* of how this experience arises does not depend on this decision. If the dualist interpretation is right, any explanation will be too weak for reducing qualitative differences to physical differences; if the physicalist is right, these explanations will finally add up to full-blown reduction of mental to

physical properties. Thus, the instability objection that Stroud raises seems to be a red herring. It just repeats the pattern of reasoning that, in section 2, we found in Fechner and James: the empirical parallelism between mental and physical phenomena needs some 'deeper' metaphysical explanation. This also seems to be the motivation when it is said that a Kantian question needs to be addressed here. It is not enough, it is argued, to explain how consciousness emerges; one also needs to explain *how such an explanation is possible* (Bieri 1996; Birnbacher 2002). But are we, as philosophers, really in a position to request such a further explanation? When scientists explain why some phenomenon occurs, they also show how such an explanation is possible: simply by providing this explanation. To claim that this could not be a 'complete' or 'full' explanation means to fall into the trap set up by the skeptic.

This brings me to the second objection that modest naturalism is a self-destructive philosophical position because it asks us to hand over all problems to the empirical sciences. There would then nothing left for philosophy to do. Clearly, anti-naturalists are here in a more comfortable position. If philosophy can pass judgment on what science can or cannot achieve, it can also carve out for itself a domain of problems that are its exclusive domain. The problem of 'fully' explaining consciousness would be a prime example of such a genuine philosophical problem.

My response to this second objection is similar to the first one. We cannot assume, without loosing the debate against the skeptic, that philosophy is able to solve deep metaphysical puzzles. This is what advocates of physicalism and dualism hope for. Naturalism suggests that these puzzles should be 'translated' into problems of the empirical sciences. This requires someone to do the translation. It is not obvious how these problems should be addressed by the different methods and conceptual frameworks used in different sciences. Philosophy can contribute to this project by showing how these differences in method and concepts may be bridged.

Compared with the tasks that metaphysically inspired philosophers set themselves, this goal of methodological and conceptual clarification seems a very minor one. In some sense this is true, and it is part of the modesty of naturalism. In another sense, however, naturalism is not a modest view at all. It claims that the skeptics are mistaken and that consciousness is not

something deeply mysterious. Without the naturalist, we would not be able to uncover the erroneous moves from which skeptical arguments receive their alleged power.[*]

8. REFERENCES

Almeder, R. 1998: *Harmless Naturalism. The Limits of Science and the Nature of Philosophy.* La Salle: Open Court.

Armstrong, D. 1980: "Naturalism, Materialism, and First Philosophy." In: Armstrong, D. *The Nature of Mind and Other Essays.* New York: Cornell University Press, 149-165.

Bickle, J. 1998: *Psychoneural Reduction. The New Wave.* Cambridge, MA: MIT Press.

Bieri, P. 1996: "Was macht Bewusstsein zu einem Rätsel?" In: Metzinger, T. (ed.) *Bewusstsein. Beiträge aus der Gegenwarts-Philosophie.* Paderborn: Mentis, 61-77.

Birnbacher, D. 2002: "Läßt sich das Bewußtsein erklären?" In: Pauen, M. and Stephan, A. (eds.) *Phänomenales Bewusstsein – Rückkehr zur Identitätstheorie?* Paderborn: Mentis, 165-176.

Brentano, F. 1982: *Deskriptive Psychologie.* Edited by Chisholm, R. M. and Baumgartner, W. Hamburg: Meiner.

Carnap, R. 1931a: "Überwindung der Metaphysik durch logische Analyse der Sprache." *Erkenntnis* 2, 220-441.

Carnap, R. 1931b: "Die physikalische Sprache als Universalsprache der Wissenschaft." *Erkenntnis* 2, 432-465.

Chalmers, D. 1996: *The Conscious Mind.* Oxford: Oxford University Press.

Chalmers, D. 1995: "Facing up to the Hard Problem of Consciousness." *Journal of Consciousness Studies* 2, 200-19.

Craig, W. L. and Moreland, J.P. (eds.) 2000: *Naturalism. A Critical Analysis.* London: Routledge.

De Caro, M. and D. Macarthur (eds.) 2004: *Naturalism in Question.* Cambridge, MA: Harvard University Press.

Du Bois R. E. 1974: "Über die Grenzen des Naturerkennens." In: Du Bois R. E. *Vorträge über Philosophie und Gesellschaft*. Hamburg: Meiner.

Fechner, C. G. 1860: *Elemente der Psychophysik*. Leipzig: Breitkopf & Härtel.

Fink, H. 2006: "Three Sorts of Naturalism." *European Journal of Philosophy* 14, 202-21.

Heidelberger, M. 2003: "The Mind-Body Problem in the Origin of Logical Empiricism." In: P. Parrini et. al. (eds.) *Logical Empiricism. Historical and Contemporary Perspectives*. Pittsburgh: University of Pittsburgh Press, 233-262.

Hornsby, J. 1997: *Simple Mindedness. In Defense of Naive Naturalism in the Philosophy of Mind*. Cambridge, MA: Harvard University Press.

Husserl, E. 1950: *Die Idee der Phänomenologie*. Edited by Biemel, W. Husserliana II. Dordrecht: Springer.

Hylton, P. 1994: "Quine's Naturalism." In: French, P. et. al. (eds.) *Midwest Studies in Philosophy*. Notre Dame: University of Notre Dame Press, 261-282.

Keil, G. and Schnädelbach, H. (eds.) 2000: *Naturalismus. Philosophische Beiträge*. Frankfurt am Main: Suhrkamp.

Kim, J. 2003: "Logical Positivism and the Mind-Body Problem." In: Parrini, P. et. al. (eds.) *Logical Empiricism. Historical and Contemporary Perspectives*. Pittsburgh: University of Pittsburgh Press, 263-278.

Kitcher, P. 1992: "The Naturalists Return." *Philosophical Review* 101, 53-114.

Levine, J. 1983: "Materialism and Qualia: the Explanatory Gap." *Pacific Philosophical Quarterly* 64, 354-361.

McDowell, J. 2004: "Naturalism in the Philosophy of Mind." In De Caro, M. and Macarthur, D. (eds.) *Naturalism in Question*. Cambridge, MA: Harvard University Press, 91-105.

McDowell, J. 1996: "Two Sorts of Naturalism." In: McDowell, J. *Mind, Value and Reality*. Cambridge, MA: Harvard University Press.

McGinn, C. 1989: "Can We Solve the Mind-Body Problem?" *Mind* 98, 349-366.

McGinn, C. 1999: *The Mysterious Flame*. Cambridge, MA: The MIT Press.

Mills, E. 1996: "Giving up on the Hard Problem of Consciousness." *Journal of Consciousness Studies* 3, 26-32.

Quine, W. V. O. 1981a: *Theories and Things*. Cambridge, MA.: Harvard University Press.

Quine, W. V. O. 1981b: "Five Milestones of Empiricsm." In Quine, W. V. O.: *Theories and Things*. Cambridge, MA: Harvard University Press.

Quine, W. V. O. 1966: "The Scope and Language of Science." In: Quine, W. V. O.: *The Ways of Paradox and Other Essays*. New York: Random House.

Quine, W. V. O. 1995: "Naturalism; or, Living within One's Means." *Dialectica* 49, 251-261.

Rey, G. 1983: "A Reason for Doubting the Existence of Consciousness." In: Schwartz, R. J. et. al. (eds.) *Consciousness and Self-Regulation. Advances in Research and Theory*. New York, London: Plenum Press, 1-39.

Sellars, R. W. 1922: *Evolutionary Naturalism*. Chicago: Open Court.

Stich, St. 1996: *Deconstructing the Mind*. Oxford: Oxford University Press.

Stroud, B. 1996: "The Charm of Naturalism." *Proceedings and Addresses of the American Philosophical Association* 70, 43-55.

Stubenberg, L. 1997: "Austria vs. Australia: Two Versions of the Identity Theory." In: Lehrer, K. and Marek, J.Ch. (eds.) *Austrian Philosophy, Past and Present*: *Essays in Honour of Rudolf Haller*. Dordrecht: Kluwer, 125-146.

Tetens, H. 1996: "Die Rettung der mentalen Phänomene? Eine kurze Regieanweisung für einen nicht-reduktiven Materialismus." In: Krämer, S. (ed.) *Bewußtsein. Philosohische Beiträge*. Frankfurt am Main: Suhrkamp, 155-166.

Van Gulick, R. 1996: "Was würde als eine Erklärung von Bewusstsein zählen?" In: Metzinger, T. (ed.) *Bewusstsein. Beiträge aus der Gegenwartsphilosophie*. Paderborn: Mentis, 79-101.

Van Gulick, R. 2001: "Reduction, Emergence and Other Recent Options on the Mind/Body Problem. A Philosophical Overview." *Journal of Consciousness Studies* 8, 1-34.

Van Gulick, R. 2002: "Nichtreduktiver Materialismus – noch immer das beste Angebot auf dem Leib-Seele Basar." In: Pauen, M. and

Stephan, A. (eds.) *Phänomenales Bewußtsein – Rückkehr zur Identitätstheorie?* Paderborn: Mentis, 297-326.

Vidoni, F. 1991: *Ignorabimus. Emil Du Bois-Reymond und die Debatte über die Grenzen wissenschaftlicher Erkenntnis im 19. Jahrhundert.* Frankfurt am Main: Peter Lang.

Wilkes, K. 1984: "Is Consciousness Important?" *British Journal of the Philosopohy of Science* 35, 223-243.

Indeterminacy of a Free Choice: Ontic, Epistemic, or Logical?

Helmut Fink
University of Erlangen-Nürnberg

1. THE PROBLEM OF ALTERNATE POSSIBILITIES IN THE DEBATE ON FREE WILL

Cognitive events are characterized by a notorious dichotomy of their possible modes of description. On the one hand, an introspective description can be given by a succession of mental states of a person. Mental states refer to conscious perceptions, thoughts and feelings the person has. These states cannot by themselves be "seen" from the outside. Strictly speaking, they can only be given meaning from a first-person perspective: It is only me who knows what it is like to be myself. On the other hand, an external description of cognitive events can be given by the dynamics of the neural states of the person's brain. Neural states are states of matter. They can be measured from the outside and get their meaning from a third-person perspective.

The question how exactly these two modes of description relate is a very far-reaching one. In philosophical tradition, this question lies in the center of what has become well-known as the mind-body-problem. In modern neuroscience, the search for neural correlates of mental states has proven very fruitful. In cognitive science as in daily life, comparing mind talk with body experience is the ultimate reason for ascribing mental states, consciousness and the like to other living beings. In any case, it is a basic building block of every naturalistic world-view (and even some moderate anti-naturalistic ones) that mental states are always accompanied by a material basis.

In this article, we are only concerned with a special case of cognitive events, albeit a very important one: the case of a free choice of an agent.

Driven by experimental results on the temporal order of neural versus mental states (Libet 2004) and on the deceivability of the feeling of control (Wegner 2002), the debate on free will has been intensified among scholars in science and philosophy in recent years and at the same time spread out in popular journals and the media. The neural correlates of mental states obey the laws of physics. Hence: Do we have free will, and if so, in which sense? Or do we have to accept, at least at a scientific level, that all we want and do is actually determined by the matter of our neurons—and not by us? Is free will an illusion, even though a useful one?

The relevance of this debate exists without doubt: Our view on acting persons in contrast to moving bodies involves a notion of free will. Our customary concept of moral responsibility depends on the possibility of a free choice. And the traditional criteria for guilt in criminal law hinge upon the freedom of the culprit's choice. One of the (perhaps very few) genuine philosophical contributions to debates like this is the search for a precise reconstruction of concepts that are phenomenologically proven (or appear to be so).

For the concept of free will, three principles or, better, three conditions have been identified. The first one is the condition of *authorship*: A free decision to act should be attributable to the agent. The person who decides should be, in some sense, the initiator of the decision. Second, a free decision should be *intelligible*, that is, it should be possible to give reasons and to understand reasons for and against a certain decision. Third, there should have been, in an appropriate sense, *alternate possibilities* for making the choice. It should have been possible to decide differently, so that counterfactual propositions ("if I had chosen otherwise ...") are meaningful.

As can easily be seen, these three conditions cannot be maintained simultaneously if each of them is interpreted in a strong sense: Assume a strong understanding of authorship. This implies a deterministic relation between cause and effect, because without underlying strong causal lines the attribution of authorship would become uncertain. But then, if determinism applies also to neural dynamics (as it has to), there is no room left for alternate possibilities in a given situation. Conversely, if there is no sufficient cause for the next link in a causal chain, such that alternate

possibilities show up e.g. in neural dynamics, then this causal gap undermines the attribution of authorship.

Further, if intelligibility is taken for granted, one is committed to the level of reason given by the decider. But deterministic causal lines for the underlying dynamics of his/her brain would certainly lead back to material conditions outside the decider's body and even before his/her birth. This kind of consequence argument highlights a tension between intelligibility and determinism. But the tension between intelligibility and indeterminism is at least as strong: Pure chance is the natural enemy of a definite reason.

As a consequence of these considerations, some or all of the three conditions given above have to be weakened. A strong concept of free choice would be inconsistent (an insight that had been well-known, by the way, to some philosophers long before brain imaging achieved its present-day publicity). Yet, there is no unique recipe how to weaken the conditions for free will. Since the naive expectation can only partially be reconstructed, there may be different consistent concepts of free will, according to different combinations or interpretations of the conditions maintained.

In order to save the intuition of free will (which is more than the freedom to act), we prefer a *compatibilistic* stance. In this context, compatibilism stands for accepting compatibility of the free will of persons with the determination of all neural processes by natural law. The resulting view on free will is perfectly consistent with naturalism, as has brilliantly been argued for (Beckermann 2005). To this elaborated position, we can add not much more than some explanatory remarks.

First, the key to accept compatibilism is to replace *whether* a decision is determined by *how* it is determined. A decision is called free, if the person who made it could judge his/her desires without coercion or compulsion. There is no freedom without deliberation. The freedom of a person's will lies in its coincidence with the person's long-term preferences and rational interests. In a sense, free decisions can be said to be determined by argument and reason. It is this process of deliberation by which a person acquires his/her free will. The resulting free decision is emphatically his/hers. People are the authors of their decisions. But there is no ultimate authorship that could substitute for a decision's causal prehistory.

Second, the key to avoid methodological confusion is to keep the different modes of description separated. One can either give reasons or study causes, but not both at the same time. To give a reason is to communicate a particular content of a person's conscious thought, while a cause has to be looked for in the behavior of matter. Reasons and causes are elements of two completely different language games: "mind talk" on the one hand and "matter talk" on the other hand. One may object that a good reason is never subjective but always intersubjective. This is true insofar as one may abstract from all individual instantiations of a mental content, and it is true in particular if this content is a reason. But nonetheless reasons, like intentions or aims, are not part of our physical description of the world; causes are.

In a causally closed material world, there is no proper causal role to play for the mind as distinct from the matter. Ontologically, one may choose from a range of positions compatible with this insight, among them e.g. neutral monism or physicalistic epiphenomenalism. We need not be so ambitious as to decide this question here. In any case, mixing up mental states with neural states is an epistemological category mistake. Reasons have no causal power just as causes have no persuasive power. Hence, one should beware of hybrid concepts like "mental causation". Causal links can only be found *within* the physical mode of description. By virtue of neural correlates of consciousness (or conscious correlates of neural dynamics), it is always possible and often useful to change the mode of description of a cognitive event. But the link *between* these modes is not a causal one.

For a reason to "determine" a decision, a causal bypass within the physical world is called for: switch from the instantiation of the reason to the underlying neural state, study its causal dynamics, and then switch back to the mental mode of description. If you regard this procedure as irrelevant to your purpose, then you are most likely interested not in causation but in reasoning. For the purpose of reasoning, it is methodologically consistent to stay in the mental mode of description all the time. It is in this mode that a free decision fulfills the condition of intelligibility.

Third, the key to understand authorship is to reconstruct the concept of a person. In addition to the *modes* of description introduced so far, there are also different *levels* of description: One can either stick to the most

elementary building blocks of a complex object (or subject) to be described, or introduce concepts at a more phenomenological level. For example, phenomenological thermodynamics is provided with its own concepts, although the underlying degrees of freedom of the matter under study can be measured and described at the microscopic level (at least in principle). This microscopic description may be regarded more fundamental, but it is not appropriate for every purpose. To be sure, a macroscopic description must not contradict the predictions on the micro-level. But the hierarchical structure of the world we live in suggests using phenomenological concepts on the macro-level without a bad methodological conscience. The concept of a person is on the macro-level. In this concept, aspects of the physical as well as of the mental mode of description are combined: A person consists of his/her body including the brain, and also of the stream of consciousness accompanying its neural dynamics. Persons are highly structured accumulations of matter, together with the perceptions, thoughts and feelings that appear from the correlated internal ("first-person") perspective.

To introduce the personal level of description in this way may not be sufficient in order to define the concept of a person. As was stressed by Frankfurt (1971), persons should be characterized by the structure of their will. But in any case, introducing the personal level is necessary to trace back authorship to the internal control mechanisms of a person. Once this personal level of description is accepted, there is no point anymore in looking for prior and exterior causes explaining the person's free decision. To be sure, the existence of such causes is not denied. However, considering them is not appropriate for the purpose of understanding authorship and is, therefore, stopped by a methodological cut.

Propositions about persons cannot simply be replaced by propositions about some of their parts. Phrases like "the brain decides" instead of "the person decides" give rise to misunderstandings concerning free will, all the more if a deterministic brain dynamics is supposed. Mixing up a whole and its parts in this way amounts to the so-called mereological fallacy. Notably, all our remarks about the concept of a person refer to methodological consistency within a naturalistic approach. They are not meant to suggest that persons are metaphysical entities or somehow excluded from nature's law. In fact, the opposite is true.

Combined, the conceptual reconstruction of "free will" involves a mental mode and a personal level of description. The concept of free will is of great phenomenological importance. To call a decision "free" helps to distinguish its mechanism from pure chance on the one hand and from coercion or compulsion on the other one. This phenomenological distinction has nothing of an illusion. However, our conceptual reconstruction seems not yet complete. Taken for granted authorship (by persons) and intelligibility (by reasons), what about the condition of alternate possibilities? How can we do justice to the strong intuition that a person's decision has been free only if they could have decided differently?

2. THREE KINDS OF INDETERMINACY

If alternate possibilities are to play some role in the description of a decision at all, some element on some stage of the decision process must be, at least in some sense, indeterminate. Conceptually, one can distinguish three different kinds of indeterminacy: ontic, epistemic, and logical. Ontic indeterminacy refers to "things as they really are", while epistemic indeterminacy refers to "things as they appear to us". Epistemology is about our knowledge and its limits. Epistemological propositions about perceptions, thoughts or feelings do not imply ontological claims about the exterior world. Both ontic and epistemic indeterminacy will be discussed in the present section. Logical indeterminacy will be shortly explained at the end of this section but criticized in the subsequent one.

It is reasonable to accept that there is ontic indeterminacy in the world. One of the two major revolutions in 20th century physics has given rise to quantum theory. Compared to classical physics, among the basic characteristics of quantum theory is its power to quantify ontic (objective) indeterminacy, e.g. by Heisenberg's uncertainty relation, and to provide probabilistic predictions for the not-yet-determined values of physical (measurable) quantities. Quantum probabilities do not merely reflect the subjective ignorance of an observer (as classical probabilities do). Leaving aside bizarre extensions of the quantum formalism that are on the run from Occam's razor (like Bohmian mechanics), one can prove mathematically that it is inconsistent to attribute definite values of all physical quantities to

a quantum system simultaneously. Definite values appear only in a measuring apparatus, where they become visible e.g. as a pointer position.

Can we make use of quantum theory in the area of free decisions? We think the answer to this question is in the negative. It is true that there is a break in the deterministic behavior of matter wherever a quantum effect shows up: If a quantum system is prepared twice in exactly the same quantum state and exactly the same quantity is measured each time, one can nevertheless obtain two different results. So there are alternate possibilities for the behavior of matter even at the most microscopic level of description. But these possibilities are of no use for the concept of free will, as was indicated quite at the beginning of this article: Breaking the chain of causal links inevitably undermines the attribution of authorship. If ontic indeterminacy comes into play on a critical stage of a decision process, then this decision is not free but arbitrary (within the given range of possibilities). A relevant part of the neural dynamics correlated with the process of deliberation is subject to pure chance then. Every quantum process is like a tiny random number generator in the decider's brain. This is a model of diminished responsibility rather than of free will.

Within the interpretation debate on quantum theory, there have been a few attempts to relate measurement outcomes with the consciousness of the observer. But every single outcome is a fact, and the quantum statistics of outcomes is determined by the physical situation. Facts as well as their statistics are part of the material world. They are not (only) in the mind of some observer. Quanta are not qualia—after all.

Conceptually, ontic indeterminacy seems to be irrelevant to reconstructing the essential features of free will. Even worse, randomness comes as a threat to authorship and intelligibility. Hence, pure chance is no chance of free will.

Fortunately, quantum effects seem to be irrelevant to neural behavior also empirically. The physiological processes relevant to perceptions, thoughts and feelings can be described in classical concepts, at least according to the vast majority of neuroscientists. There is no convincing neurophysiological evidence of quantum effects in brain dynamics so far. This observation completes our discussion of ontic indeterminacy. In order to avoid mixing up the different kinds of indeterminacy, we will assume a

classical, deterministic universe while considering epistemic and logical indeterminacy.

Epistemic indeterminacy refers to a limited state of knowledge either of some external observer, i.e. from a third-person perspective, or of the decider him-/herself, i.e. from a first-person perspective. Let us first consider the "view from the outside". If an observer wants to predict a free decision of someone else, he/she may study this person's usual behavior and look for determinants of the decision in the causal prehistory and in the environment of the decider. In view of physical law, the best the observer can do is to study the decider's neural dynamics.

Under natural circumstances, the observer's prediction will not always be successful. He/She may have overlooked one or another causal influence in the prehistory of the decision. Or his/her theoretical model of the human brain may have been too simple for a physical system as complex as this. Perhaps his/her computer power has not been sufficient. Concerning the concept of free will, none of these obstacles to a prediction of a person's decision appears as fundamental. Causal determinants can be investigated, theoretical models can be improved, computer power can be increased. In this way, alternate possibilities for the decision under study can be successively excluded until, in the limit of infinite ressources, the observer's prediction becomes unique. But one would not say that a free decision becomes less free just by sharpening an observer's prediction. Hence the condition of alternate possibilities in the debate on free will cannot convincingly be fulfilled by subjective ignorance of an observer.

We add three remarks. First, the knowledge about all the relevant causal influences is hard to achieve. From a practical point of view, unique predictions of free decisions will remain impossible for a long time. But our main concern here is conceptual rigor, not practical prospects. Therefore we refer to idealizations where it seems helpful.

Second, one may argue that the neural processes accompanying the deliberation of a free decider are perhaps algorithmically irreducible. If this were true, no observer of a decision process could exclude alternate possibilities by computational means until the process is finished. In other words, the neural process was the shortest possible solution to the prediction problem. In view of the imperfection of almost all human capacities we regard this assumption as not very plausible. It may be

conceivable as an exceptional case. But we are mainly interested in the regular case.

Third, one could prefer describing the decision situation on a macroscopic level rather than on a microscopic one. Even if a unique prediction is regarded possible, it may seem appropriate to refrain from specifying the situation completely. Situations in daily life are usually not being described by specifying the initial conditions of every potentially relevant degree of freedom. Rather, parts of the environment, including one's fellow human beings and their behavior, are treated like a variable. In such a partially specified description, a person's decision is still open to be influenced. Alternate possibilities appear.

Nevertheless, in a deterministic world, the person can decide differently only if at least one of the initial conditions is different. Hence the macroscopic level of description outlined here corresponds to a certain probability distribution of initial values in a microscopic description. In analogy with the methodological cut introduced for reconstructing personal authorship, it is an at first sight plausible approach to underpin the intuition of alternate possibilities by subsuming different deterministic micro-histories under a common description on a personal level.

However, coarse-graining the description in this way is by no means a necessary consequence of introducing the personal level: authorship can perfectly be reconstructed by reference to the internal control mechanisms of a person even if all initial conditions are exactly specified. In any case, it is legitimate to consider such a fine-grained description for theoretical reasons. From this perspective, however, alternate possibilities do appear as an artefact of coarse-grained descriptions. Thus, one might call alternate possibilities an epistemological illusion, if epistemic indeterminacy from a third-person perspective were the only approach to reconstruct their meaning.

Independently of the level of description chosen, and notwithstanding the three remarks just listed, we think that "unpredictable" is a very poor reconstruction of "free". As far as the "view from the outside" is concerned, we conclude that there is no need to identify alternate possibilities at all. In this respect we agree with Frankfurt's influential analysis (Frankfurt 1969), in which he disentangled moral responsibility

from the principle of alternate possibilities. But still, the "view from the inside" remains to be considered.

By epistemic indeterminacy from the first-person perspective we mean the ignorance of the decider about how he/she will decide. This ignorance seems to be necessary until the decision is really made. Deliberation is essential to freedom, as we argued for in section 1. But the relevant deliberation that precedes a free decision is of a particular kind: arguments for and against different possibilities have to be weighed. Without alternate possibilities, there is no decision to make. If the decider already knows the result of his/her deliberation, one would look for the moment of decision in the past and not in the future.

Thus, this lack of prior knowledge is a convincing candidate for reconstructing the origin of the robust intuition of alternate possibilities in the debate on free will. We stress that predictability is assumed to be limited only from the internal perspective now. For an external observer (with perfect physical knowledge), every neural state of the decider, and hence also the result of the decision process, may be predictable. Decoding the decider's individual reasons from his/her neural states would be a discouragingly difficult task. We repeat that reasons have to be gained by abstraction from certain mental states, which is a different epistemological category from neural states. But by virtue of mental correlates of neural states, hypothetical predictions can be made. These can be tested by comparison with the decider's comments and actions. As a result, the external observer may one day be able to give a unique, and even intelligible, prediction of the decider's next free choice.

Nevertheless, the decider refers to alternatives that he/she regards, and has to regard, as equally possible (though, perhaps, not equally probable) during his/her deliberation process. This epistemic indeterminacy from the first-person perspective seems to be *constitutive* for the freedom of a person's will. The resulting reconstruction of free will is perfectly compatible with physical determinism. Even if the decider knows that he/she lives in a deterministic world, he/she is still confronted with subjective alternate possibilities. We remark that Walde (2006) has most recently referred to this condition for free will under the notion of epistemic openness of the future ("Epistemische Offenheit der Zukunft").

Now, one may ask a far-reaching question: What happens if the decider him-/herself is going to compute the result of his/her own future decision on the basis of the deterministic dynamics of his/her neural states? Or, what amounts to the same, if an external observer communicates a unique prediction of the result before the decider's deliberation process is finished? Certainly, the presence of this information is a major stumbling block for the consideration of equally possible alternatives. Does the decider cease to be free then? Is this scenario consistent at all?

According to Popper's classical essay (Popper 1950), a deterministic system can neither predict its own future state nor receive such a prediction from another system without threatening the validity of the prediction by this very act. Hence one might think that a decider is necessarily in a state of ignorance in comparison with his/her observers.

Strengthening Popper's conclusion, Donald M. MacKay has developed the position of logical indeterminacy (MacKay 1960; 1967). According to this position, there is a relativity between the decider's and the observer's description, which is rooted in their distinct roles. The two descriptions differ from each other, but are uniquely related to each other. According to MacKay, neither of them can be claimed to be objectively true. Therefore the decider has no ignorance, no "lack" of knowledge, and hence the attribute "logical" instead of "epistemic".

Logical indeterminacy explicitly presupposes a deterministic universe. If the position of logical indeterminacy is well-founded, then freedom of the will is built into the decider's perspective with logical necessity, and can thus be deduced (and must be recognized) by every rational observer. However, we have severe doubts about the well-foundedness of this logical indeterminacy of a free choice.

3. CRITIQUE OF LOGICAL INDETERMINACY

Doubts are possible already with respect to Popper's conclusions. Popper claimed to have discovered a fundamental epistemic indeterminism pertaining to quantum and classical systems alike. He considered classical mechanical calculating and predicting machines, so-called predictors, which are thought of as realized in the physical world. Hence, unlike the Laplacean demon, predictors have finite resources and the task to predict

the behavior of finite systems with a specified degree of precision. Popper argued for the thesis that there are prediction tasks which no predictor can perform. In particular, no predictor can fully predict its own future states.

This kind of fundamental unpredictability, if it exists, leads to the consequence that a predictor cannot compute its own disturbing influence on other systems with which it interacts. Hence, this influence should be kept negligibly small. But this is impossible if the other system happens to be itself a predictor that is concerned with the first predictor, because predictors need to amplify weak influences of their object systems. Therefore, a successful predictor of predictors would have to remain outside the "society of predictors".

The essential point is Popper's claim that no predictor can predict its own future states. We think his reasoning for this thesis is worth reconsidering. Parts of it appear to us more confusing than convincing. The intuition behind this reasoning seems to rest exclusively on the consideration of a succession of preliminary predictions, of which each represents the effects of the preceding one. But could not the problem of making a prediction whose effects shall also be predicted be treated in a self-consistent way?

In classical mechanics, a measured system may be disturbed by the act of measurement. But unlike in quantum mechanics, such a disturbance may be calculated exactly in every individual case. What appears as disturbance is nothing but a particular interaction between measurement apparatus and measured system. This interaction obeys the same deterministic laws of nature as interactions within the measured system do. There is no reason why its effects should not be calculable and predictable.

The only additional difficulty in Popper's situation is the structure of self-consistency: what is to be predicted, depends itself on the prediction. But, supposing a sufficiently powerful algorithm, equations representing such a structure may be recursively solved despite Popper's no-go claim. This is at least our, of course fallible, intuition about these matters.

MacKay's reasoning starts from the essential contrast between communicating an observer's prediction to an observed decider and shielding the decider from any influence that could invalidate the prediction. If the prediction is communicated, processing the observer's message affects the decider's neural dynamics. According to MacKay, it

can never be excluded that this invalidates the prediction. Similar to Popper, MacKay is concerned only with predictions made on an empirical basis that is disturbed if the prediction is communicated to the decider. As above, we feel entitled to consider predictions that remain true if communicated, because they anticipate and include these effects already.

MacKay concludes that the observer's prediction must not be claimed to be objectively true, that is, true for everyone who knows of it and probably wants to test it, because it cannot become true for the decider. For him/her, all such predictions are logically indeterminate until he/she makes his/her choice. To be sure, external observers ("silent onlookers") may test their predictions and find it confirmed. But the role of the decider is, according to MacKay, logically different.

Several critical comments have appeared in the literature against this concept of logical indeterminacy. MacKay replied to all of them and gave the impression that he succeeded in defending his view (MacKay 1971, 73). As an attack to MacKay's "logical relativism", Watkins (1971) is still worth reading. Watkins—as a logical non-relativist—reconstructs the status of the predictions about a decider quite convincingly by use of *ceteris paribus*-clauses.

What shall we make out of it? If logical indeterminacy is refused, we are back to epistemic indeterminacy from the first-person perspective. If there are predictions about a decider that can be claimed to be true objectively, then the decider's lack of prior knowledge deserves to be called subjective ignorance. But then, what does it mean to remove this ignorance?

As a consequence of our remarks made above about self-consistent problems, we have to postulate that there may be trajectories in the space of neural states of a decider that represent a behavior consistent with calculated prior knowledge, or with communicated true predictions. This trajectory in the physical description corresponds to a stream of consciousness in the mental mode of description. We admit that the following consideration is highly speculative and far from all practical purposes. But we think it belongs to a complete conceptual reconstruction of a free choice and hence of free will.

4. TEMPORARY DELIBERATION AS A MINIMAL CONDITION FOR FREE CHOICE

To know a true prediction about one's own future decisions looks like a contradiction in itself. It is tempting to regard "making the choice" and "knowing the result of the choice" as one and the same thing. But there is a subtle distinction that helps to avoid the contradiction alluded to above. "To know (in principle)" is not the same as "to think now (consciously) of". If someone knows something, he/she can tell you about it when being asked. This does not imply, of course, that the contents of this knowledge is permanently in his/her mind. If it happens to be in mind, we call this an *instantiation* of the person's knowledge. Consciousness is extremely selective, and only a few thoughts "come to our mind" in a given period of time. Free decisions are rooted in deliberation processes, as we saw in section 1. Judging one's desires is a mental process that needs some time. Only by this process can a decider acquire a decision and learn to look at it as his/hers. It is true that alternate possibilities are in the mind's eye of the decider as long as he/she thinks about them. But this process of deliberation may have a break. In such a break, absolutely different contents may come to the decider's mind. Later on, he/she continues with deliberation.

Now, if a decider gets knowledge about the result of his/her decision, then he/she certainly stops deliberating while this knowledge is instantiated in his/her mind. But the task of conscious deliberation is not instantly fulfilled only because its result is known. At a later moment, the decider may continue deliberating. There is, at least, no logical contradiction in this succession of mental states. Admittedly, from a psychological perspective, this patchwork stream of consciousness may seem a little schizophrenic. But we are analyzing an extreme situation, so we need not be surprised by an extreme conclusion.

Summing up, the intuition of alternate possibilities is a crucial ingredient to the project of reconstructing the concept of free will. In order to make this concept consistent and fruitful, the condition of alternate possibilities has to be weakened. The relevant alternate possibilities are grounded in epistemic indeterminacy rather than ontic or logical indeterminacy.

Ontic indeterminacy implies unpredictability of individual future events. Realized in quantum effects, this sort of indeterminacy is too strong to be helpful for a free will. Logical indeterminacy does, as far as we can see, not exist at all. Epistemic indeterminacy from a third-person perspective, ubiquitous as it may be, misses the point of free will. To be unpredictable from outside is not to be free. In addition, predictability may be increased without loss of freedom (hopefully).

Epistemic indeterminacy from a first-person perspective, however, remains constitutive for free decisions, even if it has to be restricted to certain periods of deliberation. But this restriction is a burden only on deciders who prefer predicting their decisions instead of making them.*

5. REFERENCES:

Beckermann, A. 2005: "Free will in a natural order of the world." In: Nimtz, C. and Beckermann, A. (eds.) *Philosophy—Science—Scientific Philosophy*. Paderborn: Mentis.

Frankfurt, H. G. 1969: "Alternate possibilities and moral responsibility." *Journal of Philosophy* 66, 829-839.

Frankfurt, H. G. 1971: "Freedom of the will and the concept of a person." *Journal of Philosophy* 68, 5-20.

Libet, B. 2004: *Mind Time*. Harvard: Harvard University Press.

MacKay, D. M. 1960: "On the logical indeterminacy of a free choice." *Mind* 69, 31-40.

MacKay, D. M. 1967: *Freedom of Action in a Mechanistic Universe*. Cambridge: Cambridge University Press.

MacKay, D. M. 1971: "Choice in a mechanistic universe: a reply to some critics." *The British Journal for the Philosophy of Science* 22, 275-285.

MacKay, D. M. 1973: "The logical indeterminateness of human choices." *The British Journal for the Philosophy of Science* 24, 405-408.

Popper, K. R. 1950: "Indeterminism in quantum physics and in classical physics." *The British Journal for the Philosophy of Science* 1, 117-133, 173-195.

Walde, B. 2006: *Willensfreiheit und Hirnforschung. Das Freiheitsmodell des epistemischen Libertarismus*. Paderborn: Mentis.

Watkins, J. W. N. 1971: "Freedom and predictability: an amendment to MacKay." *The British Journal for the Philosophy of Science* 22, 263-275.

Wegner, D. M. 2002: *The Illusion of Conscious Will*. Cambridge, MA: MIT-Press.

What Naturalists Always Knew about Freedom: A Case Study in Narrative Sources of "Scientific Facts"

Winfried Löffler
University of Innsbruck

1. TWO POPULAR SUMMARIES

Perhaps the central debate in German-speaking popular science over the last years was on brain research and its possible impact on our account of the human being, with freedom of will as the central issue. But there is more going on than just a renaissance of the old philosophical determinism/indeterminism debate, discussions are also extending to possible consequences for our conceptions of responsibility, guilt, crime and penal law.

Among the main participants in the debate are Wolf Singer, Wolfgang Prinz, and especially Gerhard Roth, the latter presumably having the most numerous readership among a wider audience due to some paperbacks in a high-class book series. All of them defend a clearly naturalist, determinist account of man. Slogans like "We don't do what we want, but we want what we do" have become a sort of naturalist *mantra*, and the protests of philosophers from almost all kinds of schools have only lead to slightly more diplomatic theses so far. The tension between such claims and common sense is usually handled by various sorts of conventionalism, epiphenomenalism or fictionalism about concepts like self, authorship and responsibility. In Roth's case, the naturalist account is combined with a sort of radical constructivism from the beginning. It is not my task to comment on the stability of such a philosophical blend, since much has been said on that by others.

One of the curious features of the debate is that the freedom issue is partly regarded as an a priori one, partly as an empirical one by the same people. On the one hand you can read theses like the following:

> In order to find out that we are determined we would not need the Libet experiments. The idea of a free human will is in principle incompatible with scientific reasoning. Science presupposes that everything that happens has its causes and that one can find these causes. For me it is not understandable that someone who does empirical science can believe that free, i.e. non-determined action is conceivable. (Prinz 2004, 22)[1]

On the other hand the same people put together ample empirical material that they consider to be evidence for determinism, and the reference to empirical findings is surely the decisive argument for the public reception and credit of these claims. Among these findings are of course the Libet experiments (in their refined form conducted by Haggard & Eimer),[2] results of the social psychologists Wegner & Wheatley, results of Brasil-Neto, Pascual-Leone et al. on actions under magnetic stimulation of the brain, and the early stimulation experiments of Penfield & Rasmussen and Delgado on open brains of conscious patients since the 1930s. Here is a popular nutshell-summary of these findings by Gerhard Roth:

---

[1] "Um festzustellen, dass wir determiniert sind, bräuchten wir die Libet-Experimente nicht. Die Idee eines freien menschlichen Willens ist mit wissenschaftlichen Überlegungen prinzipiell nicht zu vereinbaren. Wissenschaft geht davon aus, dass alles, was geschieht, seine Ursachen hat und dass man diese Ursachen finden kann. Für mich ist unverständlich, dass jemand, der empirische Wissenschaft betreibt, glauben kann, dass freies, also nichtdeterminiertes Handeln denkbar ist." (All translations W. L.)

[2] Not only for the sake of brevity, I will not address the Libet/Haggard/Eimer experiments in this paper. Recent experiments by Herrmann et al. (2005) considerably reduced their relevance. These experiments confirm the suggestion spelled out by numerous interprets in the past that the readiness potential is not more than an unspecific expectation activity of the brain, and not a determination of the action. By combining the Libet-experiment with choice-reaction task, Herrmann et al. convincingly show that the readiness potential is already present before exposition to the relevant information, i.e. at a time when the willing process cannot even have *begun*.

Test persons can subliminally (e.g. via masked stimuli) by experimental tricks, hypnosis or brain stimulation be caused to actions of which they later claim that they *willed* them (Penfield and Rasmussen, 1950; Wegner, 2002; Roth, 2003). (Roth 2004, 15; similar Roth 2006, 10)[3]

Dozens of similar summaries can be found in literature. And as they stand, they seem to provide a massive empirical backing for determinism. Even our strong feeling of authorship and control can be proven to be an illusion, so we are told, but authorship and control is traditionally regarded as one of the conditions for an ontologically respectable conception of freedom. Summaries like that find a broad audience, they are taken for granted by many people including philosophers, scientists from various disciplines, science journalists and science politicians. Sometimes such summaries even get a bit face-lifted, consciously or unconsciously. An example is the following passage from GEO, a popular science magazine with thousands of readers. In an otherwise very careful, critical and balanced article on the consequences of neuroscience, the German neuroscientist Franz Mechsner reports the state of research as follows:

In his book *Das Gehirn und seine Wirklichkeit* Gerhard Roth, professor of brain research at the University of Bremen, describes experiments which are illustrative in this point. The experiments were carried out on patients whose skulls had to be opened for medical reasons. If certain cortex areas of the brain (which is insensitive to pain) were stimulated by electrodes, e.g. an arm could be raised. Asked for the reason of their movement, the patients regularly [regelmäßig] claim to have willed them. Stimuli in deeper structures like the thalamus also caused movements. But the patients perceived them as unintentional or even against their will. (Mechsner 2003, 81, my italics)[4]

---

[3] "Man kann Versuchspersonen unterschwellig (z.B. über maskierte Reize) durch experimentelle Tricks, Hypnose oder Hirnstimulation zu Handlungen veranlassen, von denen sie später behaupten, sie hätten sie *gewollt* (Penfield and Rasmussen, 1950; Wegner, 2002; Roth, 2003)."

[4] Gerhard Roth, Professor für Hirnforschung an der Universität Bremen, beschreibt in seinem Buch *Das Gehirn und seine Wirklichkeit* Experimente, die hierzu Aufschluss geben.
"Vorgenommen wurden sie an Patienten, deren Schädel aus medizinischen Gründen geöffnet werden musste: Reizte man bei ihnen mit Elektroden am (schmerzunempfindlichen) Gehirn gewisse motorische Cortex-Areale, konnte sich

The problem here is not only that we get the wrong impression that these experiments were carried out just recently at the University of Bremen. Rather, the problem is the newly inserted word "regularly". This really leaves nothing to desire for the naturalist: It seems now that we have easily repeatable experiments with strict correlations. Mechsner's rendering of the matter is not an exception; in numerous similar texts it is suggested as commonplace that actions (i.e. behavior with an accompanying "inner side" like intentions, plans, explanations etc.) could be triggered by external stimulation of the brain.

Critical readers might become suspicious here. Beyond medico-technical problems, should it really so easily be possible to cause people to movements which they report as willed? Would not the test persons at least become suspicious after a certain number of rounds? If experts who really conduct experiments in empirical brain research are being asked about such findings, they usually answer like "… never heard. Of course you can cause various sorts of spasms, tremors, seizures, emotional outbursts, inhibitions, even movements of limbs by stimulation, but never *actions*. Test persons always report that these effects somehow came from outside, for example that they can't resist to a strange desire to move the arm, but in any case that these movements are not willed by them."[5]

2. SCOPE, CONSTRAINTS AND DISCLAIMERS

The thesis of my paper is that these seemingly robust empirical claims, as they are boasted by Roth and others, are flatly wrong. According to my investigation in the history of the alleged research, there are no empirical results showing that full-blown actions (i.e. behavior with an accompanying phenomenological appearance like intentions, a feeling of

---

etwa ein Arm heben. Nach dem Grund ihrer Bewegung gefragt, behaupteten die Betroffenen regelmäßig, sie gewollt zu haben. Reize in tiefer liegenden Strukturen wie dem Thalamus lösten ebenfalls Bewegungen aus. Doch die Patienten empfanden sie als unbeabsichtigt oder sogar gegen ihren Willen zustande gekommen."

[5] For a summary of the literature see, e.g., Halgren and Chauvel 1993. Nothing of the material summarized here points to the direction of a stimulation of something like actions.

control and authorship) can be caused by brain stimulation and similar techniques as described by Roth and others.

This of course raises the question how such bold claims can emerge almost *ex nihilo* in the literature. I intend to show that this piece of neuromythology was created by a mixture of sloppy citations, confidence to hearsay, over-interpretations, confabulations, slight mistranslations, and confusions of probabilistic and strict correlations. Over the years, these mistakes seem to have established a narrative tradition dense enough to substitute empirical findings. Some naturalists obviously have always known what empirical research could only convey.

In order not to be misunderstood, some constraints and disclaimers on my agenda seem appropriate.

Firstly, the scope of this paper is in fact tiny—it is not more than a case-study. My question is only whether this particular, aforementioned claim that full-blown actions with the feeling of authorship can be caused by external stimulation is empirically warranted. Though my answer here will be to the negative; I do of course not doubt that there is a mass of evidence that actions, decisions and perceptions can be influenced and biased in countless ways.

Secondly, my claim is a purely factual one, not an "in principle" one. I only show that the purported results from the past do not prove what they are supposed to prove. I do not exclude that someone at some time could perhaps really design an experiment where it is plausible that full-blown actions can be triggered.

Thirdly, I do not aim at defending any particular account of human freedom, especially not an incompatibilist one. I just scrutinize the empirical backing of some claims.

Lastly, I do not want to promote any postmodernist ideas (of science as a whole as narrative, etc.). When talking about narrative, I mean it in the straightforward, all-day sense and not in the sense of Lyotard and others. But I found no better word to label the astonishing development that will henceforth be described.

3. DISENTANGLING PROBABILISTIC AND STRICT
CORRELATIONS

Let us start with a look at one of the more elaborate and detailed summaries that Gerhard Roth offers about earlier research:

> Electrical stimulations of the cortex were amply conducted by the Canadian neurologist Wilder Penfield since the 1930s. […—here follows a closer description of the epilepsy patients, W. L.]. Stimulation in points of the somatosensory cortex directly in front of the central fissure lead—depending on the place—to a tingling in certain parts of the body, stimulation of the primary motor cortex to spasms of particular muscles or muscle groups, stimulation of the premotor and supplementary motor cortex to complete movements of limbs (Penfield 1958). The patients reported they could not resist these movements, they perceived them as "forced upon them". Conversely, under stimulation of certain areas in these premotor areas they were not able to execute movements they *wanted* to execute, i.e. cortex stimulation lead to an inhibition. In a number of cases, however, stimulation of a cortex area near the foot of the central fissure at the border to the Sylvic fissure reliably lead to the will resp. desire to move the left resp. right hand or the left or right foot (Penfield and Rasmussen 1950).

> The Spanish neurologist José Delgado reported that under similar conditions as in Penfield stimulation of the rostral part of the so-called internal capsule [i.e., …] lead to movements of the patient which he ascribed to himself. Similarly, by transcranial magnetic stimulation (TMS) the neurologist Brasil-Neto could cause finger movements which the test person described as "willed" (both results cited after Wegner 2002). (Roth 2003, 515f)[6]

---

[6] "Elektrische Reizungen der Hirnrinde wurden extensiv vom kanadischen Neurologen Wilder Penfield seit den dreißiger Jahren des vorigen Jahrhunderts durchgeführt [... – hier folgt eine nähere Beschreibung der Epilepsiepatienten]. Eine punktuelle Reizung des somatosensorischen Cortex direkt vor der Zentralfurche führte je nach Ort zu einem Kribbeln in bestimmten Körperteilen, eine Reizung des primären motorischen Cortex zu Zuckungen einzelner Muskeln oder Muskelgruppen, eine Reizung des prämotorischen und supplementärmotorischen Cortex zu kompletten Bewegungen von Gliedmaßen (Penfield, 1958). Die Patienten berichteten dabei, sie könnten diesen Bewegungen nicht widerstehen, sie kämen ihnen ‚aufgezwungen' vor. Umgekehrt waren sie bei Reizungen bestimmter Areale in diesen prämotorischen Arealen nicht in der Lage, Bewegungen auszuführen, die sie ausführen wollten, d.h. die Cortexstimulation übte eine Hemmung aus. Bei einer

We see that Roth refers to three groups of findings (by Penfield & Rasmussen, Delgado, Brasil-Neto), and as a bundle they apparently make a strong case for the possibility to stimulate full-blown actions. All of them sound like strict correlations between stimulation and action. But a closer look reveals that they are not all of that same kind: the last-mentioned experiment by Brasil-Neto, Pascual-Leone and others (Brasil-Neto et. al. 1992) only conveyed a weak probabilistic correlation. The experiment ran as follows: Test persons were instructed to arbitrarily move either the left or right finger. When their motor cortex was stimulated by magnetic pulses on the left or right hemisphere, they moved the opposite finger somewhat more frequently, although they subjectively believed in a free choice. This probabilistic dependence was only present when the movement took place within 200 milliseconds after the pulse, it disappeared at later movements. Hence Roth's description that "Brasil-Neto could cause finger movements which the test person described as 'willed'" is wrong: the general order to move came from the researchers, only the time of movement was at the test-person's choice, just some property of the movements was probabilistically influenced by the stimulation to a small extent. No actions were caused at all, and the feeling of control was only deceived in respect of the probability of left and right.

A similar comment applies to the experiments of Daniel Wegner and Thalia Wheatley (Wegner & Wheatley 1999), two social psychologists whose results are also often used by Roth and others (although not here in this particular summary). The point here is again a purported illusion of control, but the test-persons' feeling of control was only deceived about the percentage of their share in the common action of two people. The design

---

Reihe von Patienten führte jedoch die Stimulation eines Cortexareals am Fuß der Zentralfurche im Übergang zur Sylvischen Furche zuverlässig zum Willen bzw. Bedürfnis, die linke bzw. rechte Hand oder den linken oder den rechten Fuß zu bewegen." (Penfield und Rasmussen, 1950)

"Der spanische Neurologe José Delgado berichtete, dass unter ähnlichen Bedingungen wie bei Penfield die Stimulation des rostralen Anteils der so genannten internen Kapsel (d.h. ...) zu Bewegungen des Patienten führte, die er sich selbst zuschrieb. Ähnlich konnte mithilfe der Transkranialen Magnetstimulation (TMS) der Neurologe Brasil-Neto Fingerbewegungen auslösen, die die Versuchsperson als ‚gewollt' beschrieb (beide Befunde zitiert nach Wegner, 2002.)"

of the experiment is somewhat complicated: Two test-persons operating something like a two-handed computer mouse were instructed to draw approximate circles on a screen which was full with pictures of objects. Every half minute they had to bring the cursor to a stop without communicating about the place to stop. Afterwards, the persons had to judge on a percentage scale whether they had rather *intended* or just *allowed* the stop just here. As a modest distraction, the test persons heard unconnected words via headphones. In fact, one of the test persons was a confederate of the researchers. Between un-manipulated rounds, this person got the headphone command to move the cursor to a certain picture following a count-down. Hence, the stops in these rounds were primarily the effect of the confederate. Nevertheless, the real test person perceived these stops as effects of "his" action at an unduly high percentage. The percentage was especially high when the noun corresponding to the stopping-place object had recently been heard via headphone. Hence, the experiment shows that one can induce illusions about control and authorship which are—at least gradually—incorrect.[7]

No doubt, both results are interesting, but they are not groundbreaking news. That people can be manipulated in their freely chosen actions by chemical, linguistic and other means, that they can even be gradually deceived about their authorship, all that has been familiar since millennia, and whole industries live from that. (Wegner & Wheatley admit that low-budget variants of such experiments can be carried out with a bowl of salted peanuts beside your TV chair). But the results discussed so far cannot be described as cases where test persons are determined to perform actions which they wrongly attribute to themselves. The experiments by

---

[7] Wegner and Wheatley (loc. cit. 488f.) themselves admit some methodological problems concerning the experiment. The number of successful manipulated rounds is rather low (27-40 responses from 51 participants were valid at each of the four time-points checked, and only eight participants had valid responses across all four trials). The reason is that it was sometimes difficult or impossible to move the cursor to the desired stopping-place. One might also worry whether test persons after a number of manipulated rounds do not become suspicious about a possible bias. Another problem may lurk behind the fact that the manipulated rounds were inserted after a number of rounds where the stopping-decision was completely left to the real test persons. This might cause a general over-estimation of their personal share in the common action.

Brasil-Neto, Pascual-Leone *et. al.* and Wegner & Wheatley do not provide the empirical basis for the bold claims in the summaries mentioned at the beginning.

4. A MORE PROMISING EMPIRICAL BASIS? PENFIELD'S & RASMUSSEN'S "OPEN HEAD" EXPERIMENTS AND DELGADO'S PATIENT

Let us consequently turn to the other two results invoked by Roth: the old findings by the pioneers of neuroscience Penfield & and Rasmussen and by Delgado, dating back to the 1930s to 70s, when experiments at the open skull with conscious patients faced less bioethical worries than today. Penfield and Rasmussen (1950) found out that electrical stimulation of certain points of the cortex lead to various forms of tingling, spasms, emotions, movements or a felt strange desire in the limb to move. But the patients always described these effects and desires as coming from the outside, or as being forced upon them. Here are the two most interesting cases:

> CASE 7. […] A further unexpected response was that at [point] 23, on the border of the fissure of Sylvius. When this point was being stimulated, she said she felt as though she wanted to move her left hand. To verify this sensation, the operator tried to "trick" the patient by warning her that he was stimulating when he did not so. This produced no such desire. He then warned her similarly when he did stimulate. She then reported the same desire to move her left hand. […]

> CASE 8. […] When H. was stimulated, he hesitated; then he said, "My hand wants to tremble a little." He referred to his right hand (ipsilateral). The hand did tremble and continued a little time after stimulation was withdrawn, but he stopped the trembling voluntarily.

> [From the explanation to Fig. 68 on Case 8]: Stimulation at [point] H produced desire to move right hand. (W. Penfield / T. Rasmussen, *The Cerebral Cortex of Man* (1950), 120-122)

Notice the constructions "she felt as though she wanted to move her left hand" and "she reports the desire to move her left hand"; we shall come back to them later on. It is more than clear that "reporting a desire" to

move a limb is not the same as "having the intention or the wish" to move it. What Penfield and Rasmussen caused by stimulation is obviously not an action in the described, full-blown sense, but rather a strange feeling as if one's limb wanted to move.

The other source indirectly cited by Roth are the electrode experiments by the Hispano-American neurologist José M. R. Delgado,[8] also dating back to the 1950s to 70s. Delgado first summarizes a mass of experiments yielding similar results to Penfield and Rasmussen: externally stimulated tinglings and other feelings, tremblings, movements, inhibitions of movements and the like (114f). What follows then is a little note on one patient, whose case is the starting point for an astonishing example of scientific hearsay, as we shall see. Here is Delgado's original text from his book *Physical Control of the Mind. Toward a Psychocivilized Society (1969):*

> In contrast to these effects, electric stimulation of the brain may evoke more elaborate responses. For example, in one of our patients, electrical stimulation of the rostral part of the internal capsule produced head turning and slow displacement of the body to either side with a well-oriented and apparently normal sequence, as if the patient were looking for something. This stimulation was repeated six times on two different days with comparable results. The interesting fact was that the patient considered the evoked activity spontaneous and always offered a reasonable explanation for it. When asked "What are you doing?" the answers were, "I am looking for my slippers," "I heard a noise," "I am restless," and "I was looking under the bed." (Delgado, 115f.)

Notice that Delgado himself gave a very cautious and unspectacular interpretation of these observations and their relevance. He comments on the scene as follows:

> In this case it was difficult to ascertain whether the stimulation had evoked a movement which the patient tried to justify, or if a hallucination had been elicited which subsequently induced the patient to move and to explore the surroundings. (loc. cit. 116)

In private correspondence (March and April 2007) Delgado told me that he still considered these attempts of an interpretation as correct, and showed a

---

8   On Delgado's life and works see Horgan 2005.

preference for the first one: the stimulation evoked a movement which the patient could not integrate, and the patient tried to give some *ex-post-*explanation for it. This phenomenon was repeatable, but (as the text in his book had already indicated) the content of these explanations differed between the rounds.[9] That means, Delgado's patient was apparently a case of the familiar phenomenon of rationalization and not a case of an external stimulation of an action.

As Delgado confirmed to me in private communication (10th April, 2007), the note in the 1969 book is the only appearance of this patient in his numerous publications.[10] This provides further evidence that the case of

---

[9] "Repetition of ESB [=electrical stimulation of the brain, W. L.] showed that the evoked behavior was reliable but the patient gave different explanations for the movement which was not in his usual repertoire. He did not say that he had initiated the movement for a purpose: he tried to explain it 'after the fact.'" (J. M. R. Delgado, personal communication, 10th April, 2007).

[10] The bibliography of the book lists 21 articles with Delgado as principal author, and 10 with him as a co-author. I retrieved and checked all these 21 articles (and some additional ones with potentially relevant titles), but none of them documents Delgado's patient or similar cases. All these articles just cover medical and technical aspects of electrode implantation and stimulation, or lengthy rows of experiments with monkeys and cats, or they provide data about the various sorts of stimulation effects we already know. As an illustration I summarize the content of the six articles with the most promising titles: *Behavioral Changes During Intracerebral Electrical Stimulation* (Higgins, Mahl, Delgado and Hamlin 1956) reports *déjà vu* phenomena and various changes in perception and verbal and bodily behavior which took place when the brain of an 11-year-old psychomotor epileptic with previous lobotomy was stimulated. These forms of behavior seem rather complex (yet highly irrational), but the boy provides no case similar to our patient. *Emotional Behavior in Animals and Humans* (Delgado 1960) reports stimulated changes in verbal and emotional behavior, movements and *déjà vu* phenomena, but nothing like stimulated actions. *Effect of Brain Stimulation on Task-Free Situations* (Delgado 1963, listed as "in press") reports experiments with Rhesus monkeys. *Psychological Responses in the Human to Intracerebral Electrical Stimulation* (Mahl, Rothenberg, Delgado and Hamlin 1964) reports how stimulation lead to linguistic and ideational effects in one patient with intractable psychomotor epilepsy. *Free Behavior and Brain Stimulation* (Delgado 1964) is a 100 pages summary about stimulation experiments with monkeys; the wording "free behavior" just refers to the fact that the stimulation could now be done via radio-control and not—as previously—with wires that restrained the free mobility of the animals; *Intracerebral Radio Stimulation and Recording in Completely Free Patients*

his patient had by far not the importance that was ascribed to him in the subsequent narrative chain. If there had really been something like a stimulation of a free action, such a sensational result would surely have deserved an appropriate publication.[11]

5. FROM MOLE-HILLS TO MOUNTAINS: HOW NARRATIVE INFLATION WORKS5.1. The evidence so far

Let me summarize what we have found out as the empirical basis for the claims in question: There is

---

(Delgado et al. 1968) reports the application of this new technology (i.e. radio-controlled stimulation and EEG recording) for the clinical treatment of four psychomotor epileptics. Assaulting behavior reminiscent of earlier outbursts could be elicited by stimulation of the amygdala, but there is no evidence for the stimulation of "actions" either.

[11] It is also illuminating to compare the later career of our passage from *Physical Control of the Mind. Toward a Psychocivilized Society*" with its original place and character. The full text of this book is available on the internet and can easily be retrieved via search-engines, but it is rewarding to hold a paper copy of it really in one's hand and to inspect it. It turns out as a paperback for a wider audience from the 1960s multi-disciplinary book series *World Perspectives* (other volumes in the series were, e.g., Jacques Maritain's *Approaches to God*, Werner Heisenberg's *Across the Frontiers* and Ivan Illich's *Deschooling Society*). The book contains a popular overview of contemporary brain research, especially under the respect of the possibilities of controlling and influencing socially problematic behaviour. In retrospect, we might perhaps not share Delgado's unbroken optimism in this point today, some commentators even ascribed a somewhat evangelical tone to the book (see Horgan 2005 and the critical literature mentioned there), but in any case it is an interesting document of its time, the history of neuroscience and its public perception. Although designed for a wider audience, the book also contains an extensive bibliography of approx. 240 research papers, some of them with titles which are *prima facie* promising for our issue (see my footnote 10). This appearance may perhaps have lead Wegner to overestimate the importance of the aforementioned case in his book *The Illusion of Conscious Will* (see chapter 5.2 below). We may speculate that Wegner was confident that a proper documentation of the patient could easily be found in one of Delgado's 31 listed papers. In Gerhard Roth's text again, where Delgado is only indirectly cited via Wegner's book, all of this prehistory is completely concealed. From Roth's text alone, the reader gets the impression of a robust, well-documented state of research.

(1) fairly good evidence for some slight and gradual deceptions about control and authorship, which however cannot be described as external determinations to actions; moreover, we have

(2) good evidence for replicable stimulations to movements and desires to move which, however, are perceived as "forced upon" by the patients. And we have

(3) one single case of a seeming stimulation of an action which is not considered as very relevant by the researcher himself. The case is not documented in research papers but only mentioned in a book for a wider audience.

I know of no other evidence which could be interpreted as the external stimulation of an action. How can the way from this poor empirical basis to the bold claims cited at the beginning be reconstructed? How can one make a mountain from a mole-hill?

5.2. Wegner's creation of the "feeling of doing"

A hub of the recent debate is Daniel Wegner's 2002 book *The illusion of conscious will*. Wegner collects and evaluates a variety of arguments which seem to point against free will. As Roth himself admits, this book is also his source on Delgado and Brasil-Neto.

  Here is Wegner's report on Delgado's patient. Having summarized Penfield's research, Wegner comments and proceeds as follows:

> […] The movements Penfield stimulated in the brain were smooth movements involving coordinated sequences of the operation of multiple muscles, which looked to have the character of voluntary actions, at least from the outside (Penfield and Welch 1951; Porter and Lemon 1993). They just didn't feel consciously willed to the patient who did them. In this case, then, the stimulation appears not to have yielded any experience of conscious will and instead merely prompted the occurrence of voluntary-appearing actions.

> Penfield's remarkable set of observations are strikingly in counterpoint, though, with those of another brain stimulation researcher, José Delgado (1969). Delgado's techniques also stimulated the brain to produce movement, but in that case movement that was accompanied by a feeling of doing. Delgado (1969) reported,

> In one of our patients, electrical stimulation of the rostral part of the internal capsule produced head turning and slow displacement of the body to either side with a well-oriented and apparently normal sequence, as if the patient were looking for something. This stimulation was repeated six times on two different days with comparable results. The interesting fact was that the patient considered the evoked activity spontaneous and always offered a reasonable explanation for it. When asked "What are you doing?" the answers were, "I am looking for my slippers," "I heard a noise," "I am restless," and "I was looking under the bed." (Delgado, 115-116)

Wegner continues his comment as follows:

> This observation suggests, at first glance, that there is indeed a part of the brain that yields consciously willed action when it is electrically stimulated. However, the patient's quick inventions of purposes sound suspiciously like confabulations, convenient stories made up to fit the moment. The development of an experience of will may even have arisen in this case from the stimulation of a whole action-producing scenario in the person's experience. In Delgado's words, "In this case it was difficult to ascertain whether the stimulation had evoked a movement which the patient tried to justify, or if an hallucination had been elicited which subsequently induced the patient to move and to explore the surroundings (1969, 116). (Wegner 2002, 45-47)

Wegner's rendering of Penfield's and Delgado's findings is basically correct, and especially it reflects Delgado's cautious interpretation of the behavior of his patient. This interpretation is not only repeated in a literal quotation, it is even underlined by Wegner's subsequent commentary. (Wegner's summary of Brasil-Neto's magnetic stimulation experiments—which I skip here for brevity—is also correct.) Problematic, however, is Wegner's introductory remark on Delgado which may direct the readers into a certain way of looking at things. Firstly, it is misleading to say that there is a "striking counterpoint" between Penfield and Delgado (this is not the case according to Wegner's own subsequent interpretation!), and secondly, the announcement that here we have a "movement that was accompanied by a feeling of doing" is a biased interpretation not warranted by Delgado's original text. At most one could perhaps say that the patient gave *ex-post*-rationalizations of his movements, or—to modify Wegner's words—he made "movements followed by a feeling of having done."

The most important thing that has changed by Wegner's compilation is the context: As its title suggests, Wegner's book is something like a list of pro-determinist arguments, and it connects arguments of very different kinds. For instance, it is Wegner who creates the bundle Penfield/Delgado/Brasil-Neto that will uncritically be taken on by Roth. The hasty reader of Wegner's text may overlook the difference, e.g., between probabilistic and strict dependencies, and in effect the empirical case for action-stimulation may look much stronger than it actually is. However, we should not blame Wegner for that. The critical reader can still keep things apart—*if he wants*.

A last problematic point, yet one of minor importance, is Wegner's lifting of *Physical Control of the Mind* into the rank of an empirical source-book. Reading Wegner, even the critical reader may now confidently believe that Delgado's patient—may he be important or not— is at least a well-documented case. We shall see that all these problematic points will reappear in Roth's account of the issue.

5.3. Roth's creation of the "will to move"

The next and crucial step of obfuscation is done by Roth himself. Let us first compare Penfield & Rasmussen 1950 with Roth 2003. Remember the constructions "she felt as though she wanted to move her left hand" and "she reports the desire to move her left hand" by Penfield and Rasmussen. In his own rendering of these results, Roth inserts two words which completely change the meaning (italics W. L.):

> In a number of cases, however, stimulation of a cortex area near the foot of the central fissure at the border to the Sylvic fissure reliably lead to the *will resp.* desire [zum *Willen bzw.* Bedürfnis] to move the left resp. right hand or the left or right foot (Penfield and Rasmussen 1950). (Roth 2003, for the German original see footnote 6)

As we said before, "reporting a desire to move" is clearly not the same as "having the will to move", but Roth's mistranslation turns the meaning of the text in that direction. A similar observation can be made concerning Roth's use of Delgado's patient. Changing the overall message of the text into its opposite is especially easy here, namely by simply cutting away Delgado's and Wegner's skeptical postscripts. Roth also changes the

construction so that the fact that it is only *one* patient is concealed: "the patient" now appears as an abstraction (the patient in general!), and not as a reference to one particular person as in Delgado's text above. (Roth's construction "in a number of cases …" shortly before may further foster this wrong impression). And finally, if the description of Brasil-Neto's probabilistic results (false as it is anyway!) is placed immediately after the incomplete description of Delgado's patient, the reader gets the completely false impression that magnetic transcranial stimulation works as reliably as electric stimulation of the brain.

5.4. A synopsis of the textual changes

To get a synopsis of the textual changes, let us finally have a second look at Roth's core text about the empirical findings backing his claim, this time equipped with more background knowledge. The reader is invited to read the text twice, once as it stands for itself, and once including my comments (in *<italics>* and reduced type size) which mark the places where the earlier textual tradition underwent important changes.

> "In a number of cases, however, stimulation of a cortex area near the foot of the central fissure at the border to the Sylvic fissure reliably lead to the will resp. desire *<mistranslation, unwarranted insertion of "will resp."!>* to move the left resp. right hand or the left or right foot (Penfield and Rasmussen 1950).
>
> The Spanish neurologist José Delgado reported that under similar conditions as in Penfield stimulation of the rostral part of the so-called internal capsule [i.e., …] lead to movements of the *<"the" suggests generality!>* patient which he ascribed to himself. *<Delgado's & Wegner's skeptical postscripts on the patient are omitted!>* Similarly *<conceals the difference between strict and probabilistic correlations!>*, by transcranial magnetic stimulation the neurologist Brasil-Neto could cause finger movements which the test person described as "willed" (both results cited after Wegner 2002)." (Roth 2003, 516; for the German original see footnote 6.)

This strikingly inadequate use of the actual evidence might perhaps raise the suspicion of deliberate manipulation. But we should be hesitant with such a verdict; a massively biased look at evidence known from second hand, based on a firm conviction what data could only be expected, is probably the better explanation.

6. CONCLUSION

We may conclude that Roth's claim that actions (in the full-blown, phenomenologically rich sense) can be triggered by external stimulation, is not warranted, at least not by the evidence he refers to. This might suggest a more general lesson. At the beginning I mentioned the question whether the determinism problem is an *a priori* matter or can be solved on empirical grounds. I deliberately left this question open at that point. But a partial answer can be given in any case: it is surely not fruitful to treat it with *false* empirical premises.[*]

7. REFERENCES

Brasil-Neto, J. P., Pascual-Leone, A., Valls-Solé, J., Cohen, L. G. and Hallett, M. 1992: "Focal Transcranial Magnetic Stimulation and Response Bias in a Forced-Choice Task." *Journal of Neurology, Neurosurgery and Psychiatry* 55, 964-966.

Delgado, J. M. R. 1969: *Physical Control of the Mind. Toward a Psychocivilized Society*. New York: Harper & Row.

Halgren, E. and Chauvel, P. 1993: "Experiential Phenomena Evoked by Human Brain Electrical Stimulation." In: Devinsky, O., Barić, A. and Dogali, M. (eds.) *Electric and Magnetic Stimulation of the Brain and Spinal Cord*. New York: Raven Press, 123-140.

Herrmann, C. S., Pauen, M., Min, B.K., Busch, N.A. and Rieger, J. W. 2005: "Eine neue Interpretation von Libets Experimenten aus der Analyse einer Wahlreaktionsaufgabe." In: Herrmann, C. S. et al. (eds.) *Bewusstsein. Philosophie, Neurowissenschaften, Ethik*. München: W. Fink, 120-134.

Horgan, J. 2005: "The forgotten era of Brain Chips." *Scientific American* 10/2005, 66-73.

Mechsner, F. 2003: "Wie frei ist unser Wille?" *GEO* 1/2003, 65-84.

Penfield, W. and Rasmussen, T. 1950: *The Cerebral Cortex of Man*. New York: Macmillan.

---

[*] My cordial thanks to Caroline and José M.R. Delgado for patiently answering my questions.

Prinz, W. 2004: "Der Mensch ist nicht frei. Ein Gespräch." In: Geyer, C. (ed.) *Hirnforschung und Willensfreiheit. Zur Deutung der neuesten Experimente.* Frankfurt am Main: Suhrkamp, 22-26.

Roth, G. 1994: *Das Gehirn und seine Wirklichkeit.* Frankfurt am Main: Suhrkamp.

Roth, G. 2003: *Aus Sicht des Gehirns.* Frankfurt am Main: Suhrkamp.

Roth, G. [2]2003: *Fühlen, Denken, Handeln. Wie das Gehirn unser Verhalten steuert.* Frankfurt am Main: Suhrkamp.

Roth, G. 2004: "Das Problem der Willensfreiheit: Die empirischen Befunde." *Information Philosophie* 5/2004, 14-21.

Roth, G. [2]2006: "Willensfreiheit und Schuldfähigkeit aus Sicht der Hirnforschung." In: Roth, G. and Grün, K.-J. (eds.) *Das Gehirn und seine Freiheit.* Göttingen: Vandenhoeck & Ruprecht, 9-27.

Wegner, D. M 2002: *The Illusion of Conscious Will.* Cambridge, MA: MIT Press.

Wegner, D. M. and Wheatley, T. 1999: "Apparent Mental Causation. Sources of the Experience of Free Will." *American Psychologist* 54, 480-492.

PUBLICATIONS OF THE AUSTRIAN LUDWIG WITTGENSTEIN SOCIETY
NEW SERIES

Time and History presents the invited papers of the 28th International Wittgenstein Symposium 2005 in Kirchberg/W. (Austria). Renowned scientists and scholars address the issue of time from a variety of disciplinary and cross-disciplinary perspectives in four sections: philosophy of time, time in the physical sciences, time in the social and cultural sciences, temporal logic, time in history/history of time, and Wittgenstein on time. Questions discussed include general relativity and cosmology, the physical basis of the arrow of time, the linguistics of temporal expressions, temporal logic, time in the social sciences, time in culture and the arts. Outside the natural sciences, time typically appears as history and in historiography in different forms, like a history of our conceptions of time. The first chapter of the book is dedicated to the major positions in contemporary philosophy of time. Is there a real sense of past, present, and future, or is time just a special coordinate among others? What does it mean that identity persists over time? The importance of Wittgenstein for present-day philosophy notwithstanding, his ideas about time have hitherto received only little attention. The final chapter, for the first time, provides an extensive discussion of his respective views.

This wide-ranging collection of essays contains eighteen original articles by authors representing some of the most important recent work on Wittgenstein. It deals with questions pertaining to both the interpretation and application of Wittgenstein's thought and the editing of his works. Regarding the latter, it also addresses issues concerning scholarly electronic publishing. The collection is accompanied by a comprehensive introduction which lays out the content and arguments of each contribution.

Contributors: Knut Erik Tranøy, Lars Hertzberg, Georg Henrik von Wright, Marie McGinn, Cora Diamond, James Conant, David G. Stern, Eike von Savigny, P.M.S. Hacker, Hans-Johann Glock, Allan Janik, Kristóf Nyíri, Antonia Soulez, Brian McGuinness, Anthony Kenny, Joachim Schulte, Herbert Hrachovec, Cameron McEwen.

ontos
verla g

Publications of the
Austrian Ludwig Wittgenstein Society. New Series

Volume 3
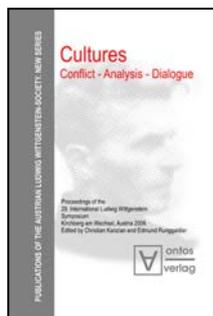
Christian Kanzian, Edmund Runggaldier (Eds.)
Cultures. Conflict - Analysis - Dialogue
Proceedings of the 29th International Ludwig Wittgenstein-Symposium in Kirchberg, Austria.
ISBN 978-3-938793-66-4
431pp., Hardcover, EUR 59,00

What can systematic philosophy contribute to come from conflict between cultures to a substantial dialogue? – This question was the general theme of the 29th international symposium of the Austrian Ludwig Wittgenstein Society in Kirchberg. Worldwide leading philosophers accepted the invitation to come to the conference, whose results are published in this volume, edited by Christian Kanzian & Edmund Runggaldier. The sections are dedicated to the philosophy of Wittgenstein, Logics and Philosophy of Language, Decision- and Action Theory, Ethical Aspects of the Intercultural Dialogue, Intercultural Dialogue, and last not least to Social Ontology. Our edition include (among others) contributions authored by Peter Hacker, Jennifer Hornsby, John Hyman, Michael Kober, Richard Rorty, Hans Rott, Gerhard Schurz, Barry Smith, Pirmin Stekeler-Weithofer, Franz Wimmer, and Kwasi Wiredu.

Volume 4
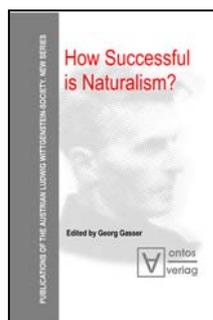
Georg Gasser (Ed.)
How Successful is Naturalism?
ISBN 13: 978-938793-67-1
ca. 300pp., Hardcover, EUR 69,00

Naturalism is the reigning creed in analytic philosophy. Naturalists claim that natural science provides a complete account of all forms of existence. According to the naturalistic credo there are no aspects of human existence which transcend methods and explanations of science. Our concepts of the self, the mind, subjectivity, human freedom or responsibility is to be defined in terms of established sciences. The aim of the present volume is to draw the balance of naturalism's success so far. Unlike other volumes it does not contain a collection of papers which unanimously reject naturalism. Naturalists and anti-naturalists alike unfold their positions discussing the success or failure of naturalistic approaches. "How successful is naturalism?" shows where the lines of agreement and disagreement between naturalists and their critics are to be located in contemporary philosophical discussion.

Volume 5
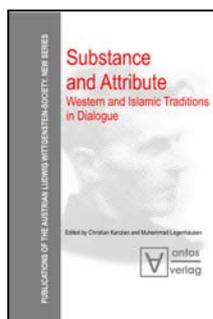
Christian Kanzian, Muhammad Legenhausen (Eds.)
Substance and Attribute
Western and Islamic Traditions in Dialogue
ISBN 13: 978-3-938793-68-8
ca. 250pp., Hardcover, EUR 69,00

The aim of this volume is to investigate the topic of *Substance and Attribute*. The way leading to this aim is a dialogue between Islamic and Western Philosophy. Our project is motivated by the observation that the historical roots of Islamic and of Western Philosophy are very similar. Thus some of the articles in this volume are dedicated to the history of philosophy, in Islamic thinking as well as in Western traditions. But the dialogue between Islamic and Western Philosophy is not only an historical issue, it has also systematic relevance for actual philosophical questions. The topic *Substance and Attribute* particularly has an important history in both traditions; and it has systematic relevance for the actual ontological debate.
The volume includes contributions (among others) by Hans Burkhardt, Hans Kraml, Muhammad Legenhausen, Michal Loux, Pedro Schmechtig, Muhammad Shomali, Erwin Tegtmeier, and Daniel von Wachter.