

épistémologie française

french epistemology

Épistémologie française/French epistemology

Ivan Vuković et Arnaud François

Institut za filozofiju Filozofskog fakulteta u Beogradu
ERRAPHIS Toulouse

Beograd 2014

Comité de rédaction/Editorial board

Giuseppe Bianco, Ondřej Švec, Ciprian Jeler
Arnaud François, Ivan Vuković

ISBN: 978-86-88803-90-8

La publication de ce volume a été soutenue par le Ministère de l'Éducation, de la science et du développement technologique de la République de Serbie, au sein du projet "Systèmes dynamiques dans la nature et dans la société: aspects philosophiques et empiriques" (n° 179041).

Le contenu

Introduction française

Introduction anglaise

I Le scepticisme au XVII^e siècle

1. Mašan Bogdanovski: "Bayle, Pyrrho and Pyrrhonism", p. 11
2. Ivan Vuković: "Epistemology and theology in Descartes' *Metaphysical Meditations*: A Cold War Reading", p. 22
3. Željko Mančić: "The Evil Genius – A History of an Argument", p. 30
4. Živan Lazović: "Cartesian Scepticism", p. 41

II Les tendances de l'épistémologie française au tournant des XIX^e et XX^e siècles : pragmatisme, holisme, intuitionnisme

5. Miroslava Trajkovski: "Léon Dumont and William James on consciousness and truth", p. 57
6. Aleksandra Zorić: "On the Duhem-Quine thesis", p. 74
7. Drago Djurić: "Pierre Duhem's Consideration of Infinity in the Middle Ages", p. 88
8. Radmila Jovanović: "Poincaré's epistemology of mathematics", p. 101

III Philosophie de la physique aux XX^e et XXI^e siècles

9. Arnaud François : "Information et entropie: Ruyer, Merleau-Ponty, Atlan, Simondon", p. 115
10. Ljiljana Dobrosavljević-Grujić : "Universalité des systèmes complexes", p. 126
11. Petar Grujić, "Scaling Down the Space", p. 135

IV Philosophie de la biologie aux XX^e et XXI^e siècles

12. Ciprian Jeler: "Normativité biologique et 'niveaux' de la vie. Remarques sur le rapport de Canguilhem à la théorie de l'évolution", p. 144
13. Eva Kamberer: "Is a Naturalistic Concept of Health Possible? Boorse's and Canguilhem's Theories of Normality", p. 161
14. Paul-Antoine Miquel : "Revenir au concept de nature", p. 175
15. Slobodan Perović: "Causes and Entities in Biology", p. 202

V Épistémologie historique

16. Ondřej Švec : "L'épistémologie française et le problème de l'objectivité scientifique", p. 222
17. Katarina Lončarević: "Foucault and Epistemology Revisited: Could Genealogy Be Seen as Political Epistemology?", p. 239
18. Philippe Huneman : "Écrire le cas – Pinel aliéniste", p. 257

Content

French introduction

English introduction

I Scepticism in the XVII century

1. Mašan Bogdanovski, "Bayle, Pyrrho and Pyrrhonism", p. 11
2. Ivan Vuković: "Epistemology and theology in Descartes' *Metaphysical Meditations*: A Cold War Reading", p. 22
3. Željko Mančić: "The Evil Genius – A History of an Argument", p. 30
4. Živan Lazović: "Cartesian Scepticism", p. 41

II Tendencies in french epistemology at the turn of the XIX and the XX centuries: pragmatism, holism, intuitionism

5. Miroslava Trajkovski: "Léon Dumont and William James on consciousness and truth", p. 57
6. Aleksandra Zorić: "On the Duhem-Quine thesis", p. 74
7. Drago Djurić: "Pierre Duhem's Consideration of Infinity in the Middle Ages", p. 88
8. Radmila Jovanović: "Poincaré's epistemology of mathematics", p. 101

III Philosophy of physics in the XX and the XXI centuries

9. Arnaud François : " Information et entropie: Ruyer, Merleau-Ponty, Atlan, Simondon", p. 115
10. Ljiljana Dobrosavljević-Grujić : " Universalité des systèmes complexes ", p. 126
11. Petar Grujić, "Scaling Down the Space", p. 135

IV Philosophy of biology in the XX and the XXI centuries

12. Ciprian Jeler : " Normativité biologique et ' niveaux ' de la vie. Remarques sur le rapport de Canguilhem à la théorie de l'évolution ", p. 144
13. Eva Kamberer: "Is a Naturalistic Concept of Health Possible? Boorse's and Canguilhem's Theories of Normality", p. 161
14. Paul-Antoine Miquel : " Individuation physique et individuation biologique : pour un nouveau concept de Nature", p. 175
15. Slobodan Perović: "Causes and Entities in Biology", p. 202

V Historical epistemology

16. Ondřej Švec : " L'épistémologie française et le problème de l'objectivité scientifique ", p. 222
17. Katarina Lončarević: "Foucault and Epistemology Revisited: Could Genealogy Be Seen as Political Epistemology?", p. 239
18. Philippe Huneman : " Une épistémologie du 'cas clinique' : l'exemple de l'aliénisme de Pinel ", p. 257

Introduction

Pourquoi aborder un thème si général, l'épistémologie, dans le cadre, plus restreint, d'une délimitation nationale ? Plus précisément, que signifie " français ", dans la locution " l'épistémologie française " ? Une tradition de lecture et de discussions (le cartésianisme) ? Une langue ? Un périmètre géographique ? Un certain nombre de pratiques institutionnelles ? Et, au-delà, une certaine méthodologie, voire certaines thèses ? Ou seulement quelques auteurs, dont la liste même est variable (Bachelard, Duhem, Canguilhem, Foucault) ? Finalement, d'où vient ce dédoublement linguistique et culturel – pourquoi la *french epistemology* ?

Partons de cette dernière question. Depuis que la philosophie analytique a commencé sa propagation rapide à travers les milieux universitaires, il y a un quart de siècle ou un peu plus, l'expression " épistémologie française " est devenue le nom d'une approche historique des sciences qui était s'était répandue en France durant le XX^e siècle, et qui était rare dans le monde anglo-saxon. Comme Ondřej Švec l'explique dans son texte, " plutôt que de résoudre le problème de l'objectivité des résultats scientifiques dans le cadre d'une traditionnelle 'théorie de la connaissance', les auteurs tels que G. Bachelard, G. Canguilhem et M. Foucault analysent les gestes concrets, les décisions et les pratiques par lesquels les différentes disciplines s'efforcent d'atteindre un savoir objectif. Les objets de la science, d'après ces auteurs, ne préexistent pas aux différentes pratiques méthodiques qui les saisissent à travers des conceptualisations toujours nouvelles, en introduisant des discontinuités dans notre façon d'objectiver le monde. "

En même temps que l'approche historique devenait la caractéristique de l'épistémologie française, l'approche de la connaissance à travers des modèles de rationalité scientifique idéalisée s'intégrait de plus en plus à la philosophie analytique et aux traditions dont celle-ci se réclame. En conséquence de cela, les différents apports d'inspiration française à l'ensemble des questionnements épistémologiques tombaient en désuétude hors de France, alors même que l'approche historique connaissait un regain d'intérêt en son propre sein.

En préparant ce volume est issu, on s'est aperçu que dans cette situation scientifique virtuellement polémique, l'attitude la plus neutre était de prendre le mot " français " en sa signification *géographique et linguistique*, et d'organiser à partir de là les positions qui se sont développés en France et en français à travers les derniers siècles (en remontant en effet jusqu'à Descartes). La nature des sujets, proposés par les auteurs eux-mêmes, a tout de suite dissipé les malentendus qui auraient pu survenir, en montrant que l'approche historique n'est qu'une des orientations épistémologiques ayant été articulées en France (à côté de l'intuitionnisme en philosophie des mathématiques de Poincaré, par exemple), que d'autres orientations figurent du reste parmi les traditions appropriées par la philosophie analytique (le scepticisme cartésien, le holisme de Duhem), que des auteurs français ont influencé les plus grands philosophes anglo-saxons (Dumont et James), ce qui d'ailleurs justifie la présence d'un certain nombre de textes en anglais dans le présent ouvrage, enfin que les auteurs français et francophones participent aux débats les plus actuels en épistémologie et en philosophie des sciences (sur des questions de physique et de biologie, abordées dans notre volume). Cela nous a permis de discuter les interprétations anglo-saxonnes des auteurs français et *vice versa* et, finalement, de transgresser les frontières linguistiques, culturelles et philosophiques que nous avions tracées d'abord.

Ce volume est issu de la collaboration, déjà ancienne, de beaucoup de ses auteurs au réseau OFFRES (Organisation Francophone pour la Formation et la Recherche Européennes en Sciences Humaines), qui a organisé à Belgrade d'abord une université d'été sur les politiques et les cultures scientifiques en juillet 2010, et puis un séminaire portant le même titre que ce volume en mai 2013. Le but principal de ces rencontres était de libérer de nouvelles sources d'inspiration et de dégager de nouvelles voies d'interrogation. Le volume présent en est le fruit.

Ivan Vuković et Arnaud François

Introduction

Why did we choose to approach such a general topic, epistemology, within national confines? What does “French” mean in this phrase “French Epistemology”? A tradition of reading and discussing (Cartesianism)? A language? A geographical area? Maybe also a certain methodology, certain theses? Or just several authors, whose list varies (Bachelard, Duhem, Canguilhem, Foucault)? Finally, where did this linguistic and cultural doubling come from – why the *French epistemology* after *l'épistémologie française*?

Let us start from this last question. Since the analytical philosophy has begun to spread rapidly through universities, a quarter of a century ago, or a little bit more, *French epistemology* became the name for the historical approach to the sciences, which was widespread in France during the XX century, and rare in the Anglo-Saxon world. Ondřej Švec explains in his paper that “rather than resolving the problem of objectivity within the confines of a traditional ‘theory of knowledge’, authors like G. Bachelard, G. Canguilhem et M. Foucault analyse concrete gestures, decisions and practices by which various disciplines try to attain objective knowledge. Objects of science, according to these authors, do not exist before various methodical practices that reach them through conceptualizations which are always new, and thereby introduce discontinuities into our way of objectifying the world.”

At the same time historicity was becoming the characteristic mark of French epistemology, the approach to sciences through idealized models of scientific rationality came to be associated with contemporary analytical philosophy and the traditions she appropriated. As a consequence, various contributions of French origin to epistemological issues in general were forgotten outside of France, while the historical approach regained much attention in France itself.

While we were preparing this volume, we realized that the neutral thing to do in this potentially polemical scientific situation would be to take the word “French” in its geographic and linguistic meaning, and to organize from there the various positions that were developed in France throughout the centuries, starting from Descartes’ time. The nature of the topics, that were proposed by the authors themselves, has quickly dissolved potential misunderstandings, showing that historicity was just one epistemological approaches that were developed in France (side by side with Poincaré’s intuitionism in the philosophy of mathematics, for example), that other orientations are among the traditions appropriated by the analytic philosophy (Cartesian scepticism, Duhem’s holism), that French authors have

influenced some of the greatest Anglo-Saxon philosophers (Dumont and James, Duhem and Quine), which justifies the presence of a number of articles written in English in this volume, and finally that French and French speaking authors participate in current debates in epistemology and philosophy of science (concerning questions of physics and biology, that were treated in this volume). This has allowed us to discuss Anglo-Saxon interpretations of French authors and *vice versa*, and, finally, to transgress linguistic, cultural and philosophical frontiers that we have drawn at the beginning.

This volume is a result of a long cooperation of many of its authors within the network OFFRES (Organisation Francophone pour la Formation et la Recherche Européennes en Sciences Humaines), which has organized in Belgrade a European summer university on politics and culture of science in July 2010, and a seminar carrying the same title as the present volume in Mai 2013. Their primary goal was to liberate new sources of inspiration and to open new paths of interrogation. This volume is the fruit of those endeavours.

Ivan Vuković and Arnaud François

Le scepticisme du XVII siècle

Bayle, Pyrrho and Pyrrhonism*

Ancient skepticism recently became quite a fashionable subject, among both historians of philosophy and epistemologists. Unfortunately, in the growing scholarship, there is little mention of Pierre Bayle. It is a bit strange, considering the fact that Bayle had a lot to say about Pyrrho and Pyrrhonism in his *Historical and Critical Dictionary*. There is almost no mention of Bayle among the scholars of ancient philosophy, because Bayle had no access to relevant sources related to the problem of establishing who the historical Pyrrho really was. The picture of Pyrrho we got from Sextus Empiricus, the only preserved original written source on Pyrrhonism, does not really correspond to the more reliable ancient second hand testimonies about Pyrrho's teaching and way of life. Sextus' writings were available to Bayle, but he had to be ignorant about the most important data that became available to us much later.

Contrary to the widespread opinion that Bayle is entirely unimportant in our reconstruction of Pyrrho's thought, Bayle exhibits a specifically French sense for the history of ideas. Some extremely interesting conjectures about who the real, living Pyrrho was, are actually extractable from Bayle's articles on Pyrrho and Pyrrhonism.¹ All the students of philosophy know that Bayle was interested in Pyrrhonism as a school of thought, but what I am about to show here is that we can learn a lot about Pyrrho himself, notwithstanding the fact that Bayle was unacquainted with the testimonies available to us today. I think we can learn a lot from the *Historical and Critical Dictionary*, which has been, unfortunately, almost totally neglected in the recent scholarship, but which is very well worth taking a look at.

It is a pity that so much more attention has been given to the anecdotal evidence.² The famous stories about Pyrrho from Diogenes Laertius abound in bizarre detail. One can learn from them that Pyrrho has always been followed by his students because he was in

* This article is the result of my research done within the project *Logico-Epistemological Foundations of Science and Metaphysics* (179067), supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

¹ Pierre Bayle, *Dictionnaire historique et critique*, Amsterdam etc., P. Brunei, P. Humbert, etc., 1740.

² The most comprehensive collection is Decleva Caizi, F. Pirrone, *Testimonianzei*, Naples, Bibliopolis, 1981.

constant danger of falling into ditches or gutters. He was under permanent students care due to the fact that he did not pay any attention to the cars, and without them putting him out of danger, he would have certainly been ran off. There is an abundance of amusing anecdotes, but in my favourite one, his teacher, Anaxarchus from Abdera, was drowning in a swamp, just when Pyrrho was passing by. Of course, Pyrrho had noticed what was happening, but he continued his walk and went away, skeptically undisturbed by the sight of his teacher in mortal danger. Eventually, Anaxarchus has made a Houdini escape from the pond and he has immediately afterwards pointed out Pyrrho to the rest of his students as a model philosopher. As a completely and perfectly tranquil person, Pyrrho did not even consider swinging over to save him.³

It is a bit strange that Anaxarchus praised Pyrrho so much, since Anaxarchus was a Democritean philosopher. This is not the only reason why the anecdotes concerning Pyrrho seem bogus, especially since we have pretty convincing evidence to the contrary. Pyrrho was so respected in his community that his fellow citizens had built a monument in his honor at the main square. Even the other philosophers in Elis were exempt from paying taxes because of Pyrrho's fame. It is quite strange that a person constantly falling into trenches and a person in constant danger of being run off by a car could have become so massively respected in his community. Another fact that makes the anecdotes suspicious is the claim that Pyrrho was at first a pupil of Bryson, son of Stilpo. We know now that Bryson was actually the father of Stilpo, and that makes the whole story highly questionable.⁴ Another suspicious narrative is the one concerning Pyrrho's alleged trip to India.⁵ Apparently, Pyrrho has never been there, but we know for sure that his teacher Anaxarchus actually has, which is probably something that motivated the widespread belief that Pyrrho visited the naked wise men of India. Furthermore, there is a superficial resemblance between Pyrrho's teaching and the Eastern ascetic philosophies.

The real question is why would anyone invent such anecdotes. In my opinion, the purpose of devising those stories was to illustrate the famous *apraxia* objection. The Stoic *apraxia* objection took a specific form in the case of Pyrrho. The objection that one cannot act unless one has at least some beliefs became a powerful weapon against the skeptics in

³ For a discussion on Diogenes' account of Pyrrhonism, see Barnes, J. 1992 „Diogenes Laertius IX, 61-166: The Philosophy of Pyrrhonism“ in H. Hase ed., *Aufstieg und Niedergang der Römischen Welt* II.36.6. Berlin, De Gruyter, 2009, pp. 4241-301.

⁴ A. Bailey, *Sextus Empiricus and Pyrrhonian Scepticism*, Oxford, Clarendon Press, 2002, p. 21.

⁵ For a detailed discussion on this issue, see E. Flintoff, „Pyrrho and India“, *Phronesis*, 25, 1980, pp. 88-108.

the hands of Stoics. The Stoics developed subtle arguments concerning the causal and conceptual connection between beliefs and action. But in the case of Pyrrho, the apraxia objection was reduced to a self-destruction objection. The anecdotal Pyrrhonean would obviously die, as a result of his all-encompassing *epochê*. That gives the apraxia objection a completely different form in the case of Pyrrho, than in the case of, for example, Arcesilaus. The anecdotes about Pyrrho just illustrate the corresponding form of the *apraxia* objection. They just show what Stoics meant when they applied their objection to his case.

Fortunately, there is a reliable source concerning Pyrrho's philosophy. It is Eusebius' *Praeparatio Evangelica*, *Preparation for the Gospel*.⁶ It is a reliable source despite everything one might think when faced with its title, indicating a work of a Christian thinker fine tuning the whole history of philosophy in order to present it as an imperfect preparation for the Biblical truth. But the Bishop of Cesarea was an extremely accurate interpreter of those he considered to be the main philosophical figures in the past. We are able to see his high reporting qualities in *Praeparatio Evangelica*, when we compare his reports with the words of philosophers whose writing we have today preserved. When it comes to Pyrrho, there is a passage in that book where Eusebius reports about Aristocles of Messene, who reports about Pyrrho's student and follower Timon, who reports about his teacher. The passage is frequently referred to as Eusebius on Aristocles on Timon on Pyrrho. It sounds like a game of broken telephones. However, Aristocles was a peripathetic philosopher who utterly disliked the skeptics, and that gives us good reasons to suppose that his report is relatively reliable, at least because Aristocles is a good source for the philosophers he hated. Let us take a good look at the Eusebius' text now.

“It is above all necessary to investigate our own knowledge; for if we by our nature know nothing, then there is no need to inquire into other matters. There have also been some among the ancients who made this pronouncement, against whom Aristotle has argued. Pyrrho of Elis was also a powerful advocate of such a position; but he himself has left nothing in writing, while his disciple Timon says that whoever wants to be happy should consider these three questions.” (*Praep. evang.* 14.18.1–5 = DC 53)

⁶ Eusebius, *Praeparatio Evangelica*, ed. E. H. Gifford, 4 vols, Oxford, Oxford University Press 1903.

Of course, this is a single, continuous passage in Eusebius' book, but I have broken it here into pieces and numbered and bulleted lists, in order to separate the main points.⁷

“Whoever wants to be happy should consider these three questions:

- First, how things are by their nature;
- second, in what way we should be disposed towards these things,
- and lastly, what will happen to those so disposed.”

This is actually a nice way to describe what was called “a way of the pyrrhonist” much later in the history of Pyrrhonism. Eusebius, in turn, gives a following report concerning Pyrrho's own description of this “way”:

“He says that he [Pyrrho] declared that”

1. “things are equally indifferntiable and unmeasurable and undecidable;”

This poses a problem, and I shall have to stop here for a while. Those familiar with Ancient Greek may see the problem immediately. The words used in the original text are *adiaphora*, *astathmeta* and *anepkrita*. They might also mean indifferent, unstable and indeterminate, because it is not clear whether they are adverbs or adjectives. Obviously, two interpretations of the text stem from the two translations, and the most influential Pyrrho scholar Richard Bett prefers the second one, which he calls the metaphysical interpretation. According to him, Pyrrho has actually meant the second triad of terms, the adjectives describing the nature of things, and not the first triad, containing the adverbs describing our cognitive shortcomings.⁸ The solution of this difficulty is impossible without a serious analysis of the next sentence from Eusebius' passage:

- “because of this neither our perceptions nor opinions tell the truth or lie.”

Eduard Zeller, the 19th century German ancient scholar and historian of philosophy has supposed that the text is wrong.⁹ It simply cannot say “Because of this...”, but “because”. It was not “*dia touto*” in the original text or in Pyrrho's teaching, but “*dia to*”. The text should go: “Because neither our perceptions nor opinions tell the truth or lie.” And it is something that allegedly makes more sense. On the other hand, Bett has held that the text is just fine, and that we should not take Greek words as adverbs, but as adjectives and

⁷ This translation is taken from S.H. Svavarsson, „Pyrrho and Early Pyrrhonism“ in *The Cambridge Companion to Ancient Scepticism*, Ed. Bett, R. Cambridge, Cambridge University Press, 2010, p. 41.

⁸ R. Bett, *Pyrrho, his Antecedents and his Legacy*, Oxford, Oxford University Press, 2000, p.16.

⁹ Eduard Zeller, *Die Philosophie der Griechen in ihrer geschichtlichen Entwicklung*, Pt. 3, Vol II, 4th edn. Leipzig, Reisch, 1909, p. 501.

treat the thesis as metaphysical, not as epistemic. This becomes a very complex issue at this point, and the ways of solving it are not particularly significant for our purposes, since I have just wanted to offer a sketch of this dilemma, in order to see how it fits into Bayle's understanding of Pyrrho.

What is really important is that in both cases we are faced with dogmatic conclusions, which is something undesirable in later Pyrrhonism. It turns out that Pyrrho was a dogmatic philosopher. He held some dogmatic beliefs either concerning the nature of things or regarding our human cognitive deficiency. That would make it difficult to explain why has Pyrrho generally been treated as the father of Pyrrhonism, the philosophical movement that had *epochê*, the suspension of judgement, at its core. The philosopher who proclaimed Pyrrho to be the founding father of Pyrrhonism was Aenesidemus. The widespread explanation for this fact is that Aenesidemus was impressed by Pyrrho's way of life.¹⁰ Falling into trenches was, apparently, the most consistent behavior with his teaching. Aenesidemus, allegedly, wanted to confront the Academic skeptics and start a schism within the skeptical movement, by instituting a new school, the Pyrrhonist school of thought. He chose Pyrrho as the mythical founder because Pyrrho was a consistent skeptic. On the other hand, the Academics, according to Aenesidemus, fell into stoicism, making the Academia a school where stoics just quarrel with stoics. In this picture, there actually was no skepticism in Academia any more. According to the same widely received explanation, Pyrrho was just a dogmatic philosopher, in the sense of negative dogmatism, claiming that knowledge is unattainable. A weird looking guy who attracted a lot of attention with his bizarre behavior and extreme theses concerning the impossibility of having knowledge about anything. In fact, he was not a Pyrrhonist at all!

Let us now turn to the second question:

2. „Because of this, then, we must not trust them,
 - but we must be without opinions and lean to neither side and remain unwavering
 - saying concerning each individual thing that it no more is than is not or both is and is not or neither is nor is not.”

¹⁰ This is almost a received opinion. For a characteristic explanation of this kind, see A. Bailey, *Sextus Empiricus and Pyrrhonian Scepticism*, Oxford, Clarendon Press, 2002, p. 72-80.

This is the most skeptical translation of the ambiguous Greek expression *ou mallon*. In later Pyrrhonism, especially with Sextus Empiricus, it becomes the heavily exploited “I determine nothing” phrase. So,

3. “What happens to those so disposed, according to Timon, is first nonassertion, and then tranquility, while Aenesidemus says pleasure”.

This is the whole passage from *Praeparatio Evangelica* and, in my opinion, the most complete and the most reliable testimony about what Pyrrho was teaching. Of course, there are significant differences between Sextus’ Pyrrhonism and Pyrrho.

However, there are points in Bayle’s interpretation of Pyrrho, developed in the relevant entries of his famous *Philosophical Dictionary*, that are very much in line both with Sextus’ picture of a Pyrrhonist and with this passage in the Eusebius’ book. Notwithstanding the fact that Bayle was not familiar with Eusebius’ writing, I shall claim that Bayle’s entries on Pyrrho and Pyrrhonism in the *Dictionary* can give us an even better insight into who he really was.

Richard Popkin was the author who did a lot for promoting Bayle. On the other hand, Popkin might be hugely wrong on many points concerning Bayle. An attempt to criticize his thoughts on Bayle would not do much harm to the seminal character of his work, and will soon give us some important exegetical results.

Bayle starts his *Dictionary* entry on Pyrrho with a peeling of an onion metaphor, which is, allegedly, Leibniz’s metaphor, at least Popkin claims so.¹¹ Being a skeptic, and that is why Leibniz was congenial with skepticism, one peels an onion until he reaches the hard core of truth. Popkin claims that Bayle’s procedure of peeling the onion ends up with nothing, no core of truth one could later rely on. When Bayle peels the last layer, nothing is left. But Popkin is fundamentally wrong because, as a self-proclaimed Pyrrhonist, Bayle explicitly proposed a never ending procedure. A never ending procedure is something that is not in line with Pyrrho’s negative dogmatism, and there must have been some reasons why Bayle thought that peeling of an onion is, in the case of Pyrrhonists, a never ending procedure. And it is not clear how one could conceive of a peeling *ad infinitum*. That conception is provided by the infinite regress of reasons, the doctrine Bayle ascribed to Pyrrho.¹² It was just a conjecture about Pyrrho’s philosophy, but a very good one. Of

¹¹ Richard Popkin, *The History of Scepticism from Savonarola to Bayle*, Oxford, Oxford University Press, 2003, p. 291.

¹² Pierre Bayle, *Historical and Critical Dictionary Selections*, ed. R.H. Popkin, Indianapolis, Bobbs-Merrill, 1965, entry „Pyrrho“, remark B.

course, Bayle has never referred to the „indifferentiable and unmeasurable and undecidable“ scheme, since for him inquiry never ends with dogmatic conclusions, including those three about our cognitive shortcomings or the metaphysical nature of things, because of the infinite regress of reasons. One must put up some reasons in order to support these negative dogmatic pronouncements, and one needs some further reasons to support those reasons, and so on *ad infinitum*. Bayle thus thought that Pyrrho was able to be a genuine skeptic because he used the strategy of the infinite regress of reasons.

It is a relatively strange idea, because the Skeptics have never thought that the infinite regress of reasons is the main source of skepticism. The infinite regress is one of Agrippa's modes and it is a source of skepticism in virtue of being one of the *tropoi*, the modes of argumentation that should have provoked *epochê*. Skepticism means investigation, inquiry, examination in Greek, and the skeptics are *zetetic*, those who examine, those who investigate and are never satisfied with any positive or negative dogmatic conclusion. Bayle certainly had that in mind, and he pondered about the ways Pyrrho could have escaped negative dogmatism. Eventually, he came to the idea that Pyrrhonian examination never stops because of the fact that there is an infinite regress of reasons. In some way Bayle was an infinitist, like Klein, thinking that it is a fact of our epistemic practice, but not thinking that there is anything wrong with it, like the infinitists do. He saw something Sextus completely missed, at least when he was thinking about Pyrrho. Namely, if there is an infinite regress of reasons, and if the equipollence, *isostheneia*, the equal strength of the opposing reasons is the central thesis of Pyrrhonism, then the infinite regress and *isostheneia* must somehow be connected.¹³

This shows the great sense Bayle had for the intellectual history and the history of ideas. He actually invented a new kind of regress: if there is a regress of reasons, then those reasons are always in danger of equipollence. One could always devise opposing reasons that are of equal power. But those other reasons generate an infinite regress of reasons that support them. Equipollence stops to be a generator of skepticism, and becomes something that actually generates Pyrrho's negative dogmatism. It was a basic tenet of Pyrrhonism that *isostheneia* generates our cognitive shortcomings. That is how the Pyrrhonist felt what the situation in philosophy was like. The equipollence thesis, the thesis that one could always devise an opposing argument with equal power, is not based on argumentation – it is based on how one perceives the history of philosophy. The *isostheneia* has just reflected

¹³ *Ibid.* See the entry “Pyrrhonism”.

the way Pyrrhonist had experienced philosophy. Bayle thought that Pyrrho was himself a person that experienced philosophy in the same way, which resulted in some kind of negative dogmatic claims either about the metaphysical nature of things or concerning our cognitive deficiencies.

Bayle ascribes something to Pyrrho that explains why he is the father of Pyrrhonism. Actually, it is the best explanation, although it is just a conjecture of Bayle's. Even better, he gives philosophical reasons, philosophical explanation why is Pyrrho the father of Pyrrhonism. And the explanation is particularly interesting because it involves an explanation of the way the infinite regress of reasons is supported by the equipollence of reasons.

Popkin wanted to conclude his section on Bayle in the *History of Scepticism from Savonarola to Bayle* with the thesis that Bayle had substituted tranquility, *ataraxia*, with faith.¹⁴ Of course, there is a textual support for this claim and Popkin was undoubtedly meticulous in reading Bayle. If there is such a substitution, it has to be a matter of human psychology. The transition from *epochê* to faith is a purely psychological event, exactly like what Sextus thought to be the transition from *epochê* to *ataraxia*. Sextus was not clear how exactly tranquility follows the suspension of judgment. He claimed that there is no argumentation involved, no reasons for *ataraxia*. In Sextus' words, tranquility follows *epochê* like a shadow.¹⁵ In his depiction of Pyrrho, Bayle offers the transition as an intellectual option, but he thinks that it is a psychologically unrealistic option. He is not a skeptical fideist in that sense, which is obvious from his entries on the Manicheans and Spinoza.¹⁶ Popkin exploited the entry on Pyrrho because he thought that the substitution would explain why Bayle was a skeptical fideist.

I do not want to enter here into discussions about the reasons faith cannot be just an escape from the situation of *epochê*, why one cannot escape into Christian faith if he gets into that disposition or frame of mind. But it is important to note that, according to Sextus, *epochê* is not the suspension of all judgments one by one, because that is impossible. *Epochê* is a kind of mental disposition. It is the skeptic's frame of mind that accompanies all of his judgments. In that way, Sextus was able to generalize his suspension of judgment. And now we are able to understand Bayle's important point on Pyrrhonism that fits

¹⁴ Popkin, *The History of Scepticism*, p. 300.

¹⁵ PH 1. 29. in Sextus Empiricus, *Purrhôneiôn Hypotypôseis (Outlines of Pyrrhonism)*, ed. H. Mutschmann, Rev. J. Mau, Leipzig, Teubner, 1958.

¹⁶ See the entries "Manicheans" and "Spinoza" in Pierre Bayle, *The Dictionary Historical and Critical*, 2nd English ed., London, Printed for J.J. and P. Knapton, etc., 1734-8.

perfectly well into the Eusebius' passage on Aristocles on Timon on Pyrrho. Pyrrho mentions the transition from the suspension of judgment to tranquility in a purely descriptive manner. Remember that he says: "What happens to those so disposed..." It is something that just happens to you. Tranquility is not something you should, normatively speaking, attain. If you get into *ataraxia*, it is just something that happens to you. It is a psychological consequence of *epochê*. That was Pyrrho's thesis and it is exactly what Bayle says, perfectly in accordance with this report from Eusebius.

There is another point in Bayle, where he diverges from Sextus. Bayle claims that the transition from equipollence to the suspension of judgment cannot be accounted for in a purely descriptive way, as a purely psychological event. Sextus tried to depict it as a psychological event in order to avoid negative dogmatism. Bayle was familiar with what Sextus wrote about it, that *epochê* spontaneously follows equipollence. For if one argues in favor of *epochê*, he immediately falls into negative dogmatism. However, Bayle reasoned that Sextus picture of the mental transition is unacceptable. If one just yields to the psychological force of *isostheneia*, one is breaking his epistemic duties.¹⁷ As the equipollence is the fate of the entire philosophical enterprise, one must, normatively speaking, endorse or accept the suspension of judgment. And that is exactly what Pyrrho says in Eusebius' report. "Because of this, then, we must not trust them...", "but we must be without opinions and lean to neither side, and unwavering." This is the point where Bayle insists that Pyrrho had to retain some degree of negative dogmatism. Sextus is aware that there is a problem for him at this point, but he attempts to avoid it by neglecting the issue of epistemic duties, and claiming that *epochê* is a purely psychological consequence of *isostheneia*. Bayle claims that Sextus' account is intolerable. It is perfectly in line with what Pyrrho maintains, that we must accept the suspension of judgment because of the equipollence.

Generally speaking, we can see now how a skeptical fideist motivation to study a thinker, namely Pyrrho, could produce some interesting and significant conclusions about who Pyrrho really was. It is a result of Bayle's specific and peculiar sense for the history of ideas. In an indirect way, Elisabeth Labrousse said in her book about Bayle that he had religion in his heart, and not in his mind.¹⁸ In fact, she did not say that explicitly about Bayle, but she has just mentioned a kind of people with such mental attitude. But it is my

¹⁷ See the remarks to the entry on Pyrrhoneans in Bayle, *Historical and Critical Dictionary*, ed. R.H. Popkin.

¹⁸ Elisabeth Labrousse, *Pierre Bayle*, The Hague: M. Nijhoff, 1963-4, vol. 1, p. 269.

impression that she had Bayle on her mind, since he belonged to that kind of persons. But it is his intellectual honesty that provided him with the picture of Pyrrho that is more accurate than the picture our current scholarship delivers to us. Because he was an ingenious philosopher, he was able to see some details in Pyrrho's thinking, which we were not able to see relying just on the unreliable historical evidence.

References

- Bailey, A. (2002) *Sextus Empiricus and Pyrrhonian Skepticism*, Oxford, Clarendon Press.
- Barnes, J. (1992) „Diogenes Laertius IX, 61-166: The Philosophy of Pyrrhonism“ in H. Hase ed., *Aufstieg und Niedergang der Römischen Welt*, II.36.6, Berlin, De Gruyter), p. 4241-301.
- Bayle, P. (1734-8) *The Dictionary Historical and Critical*, English ed. 2nd ed., London, Printed for J.J. and P. Knapton, etc.
- Bayle, P. (1740) *Dictionnaire historique et critique*, Amsterdam, etc.: P. Brunei, P. Humbert, etc.
- Bayle, P. (1965) *Historical and Critical Dictionary Selections*, ed. R.H. Popkin, Indianapolis, Bobbs-Merrill.
- Bett, R. (2000) *Pyrrho, his Antecedents and his Legacy*, Oxford, Oxford University Press.
- Bett, R. (2010) *The Cambridge Companion to Ancient Skepticism*, Cambridge, Cambridge University Press.
- Decleva Caizi, F. (1981) *Pirrone, Testimonianze*, Naples, Bibliopolis.
- Eusebius (1903), *Preparatio Evangelica*, ed. E. H. Gifford, 4 vols, Oxford, Clarendon Press.
- Flintoff, E. (1980) “Pyrrho and India“, *Phronesis*, 25.
- Hase H. ed. (1992) *Aufstieg und Niedergang der Römischen Welt*, Berlin, De Gruyter.
- Labrousse, E. (1963-4) *Pierre Bayle*, vols. 1 and 2, The Hague, M. Nijhoff.
- Popkin, R. (2003) *The History of Scepticism from Savonarola to Bayle*, Oxford, Oxford University Press.
- Sextus Empiricus (1958) *Purrhôneiôn Hypotypôseis (Outlines of Pyrrhonism)*, ed. H. Mutschmann, Rev. J. Mau, Leipzig, Teubner.

- Svavarsson, S.H. (2010) „Pyrrho and Early Pyrrhonism“ in *The Cambridge Companion to Ancient Skepticism*, ed. by R. Bett, Cambridge, Cambridge University Press.
- Zeller, E. (1909) *Die Philosophie der Griechen in ihrer geschichtlichen Entwicklung*, 4th edn., Leipzig, Reisland.

Epistemology and theology in Descartes' *Metaphysical Meditations*: A Cold War Reading*

The question of intention

I would like to discuss an old fashioned and obsolete question: what did Descartes want to accomplish by writing and publishing his *Meditations on First Philosophy* or, as the French title said, his *Metaphysical meditations*?

For several reasons, and for at least a century now, such questions about the author's intentions have been suspected of irrelevance and of laughable *naïveté*. Should not the value of a book come from its intrinsic properties, and not from the external purpose its author wished to accomplish by publishing it in its own time? And why should we bother with old stories at all - why don't we take just those parts of an old text that still seem interesting today? Is not *our* intention more important than the authors' one?

It is also a very speculative question that can hardly be answered in a scientific way, because authors themselves very often if ever know what they really want: they want one thing publicly, another one consciously but silently, and a third one unconsciously. And all those wants are determined by a long list of complicated background conditions – political, economical, technological, sexual, physical, biological, psychological, social and cultural, linguistic, conceptual, ontological, epistemological, metaphysical, scientific and so on. That is why even if that question was still important, we should not expect to find an unambiguous and sufficiently corroborated answer to it.

I believe, however, that in spite of all those reservations that I fully endorse, something could be uncovered if we follow such a line. What is to be revealed is the reality of human struggle for self-preservation and for ideas. Whom and how a man had to fight,

* This article is the result of my research done within the project *Logico-Epistemological Foundations of Science and Metaphysics* (number 179067), supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

and with whom he had to take sides in order to survive and to spread his ideas? What strategy did he use? What did he fear? What did he cherish above everything else?

Such investigations bring neither scientific result nor moral conclusion, but they do bring a bit of wisdom – they can shed some light on the constraints of the life that we have chosen to live, on the roles that we might be asked or forced to play, and on different ways we feel about them.

The question of sincerity

Personally, however, I did not approach this question searching for wisdom. It came to me quite naturally and unexpectedly, when I went through the whole of Descartes papers in order to write a postface for the *Meditations* I had recently translated into Serbian. In such a text, one is supposed to give a brief overview of the author's life, time and achievements, and to say something about the book itself within those contexts. However, the first thing that struck me while I was mapping the context of the *Meditations* was a stark contrast between the way the book was presented to the readers, and the rest of Descartes life and work.

On one hand, there was this religious subtitle of the *Meditations*, *in which God's Existence and the Immortality of the Soul are Demonstrated*, and the dedication to the orthodox clergy at Sorbonne, that implied adherence to a particular kind of Christianity.

On the other hand, Descartes had left France at the age of twenty two by his own choice, and had spent the rest of his life in a Protestant country, out of the Catholic Church's reach, in a kind of carefully preserved seclusion that was supposed to bring him peace and freedom to investigate things as they pleased him.

More importantly, there was nothing Christian at all in the rest of Descartes' writings, since he had spent most of his time thinking about light and matter in all their forms, about mathematical tools for their analysis and their epistemological and ontological presuppositions.

Descartes previously wrote of God's existence and of the immortality of the soul only once, in his *Discourse on method* published seven years earlier (1637), where he proposed some of the arguments he would later repeat in the *Meditations*. It is true that he occasionally praised the Lord in his private correspondence, but he always did it in a

conventional manner, and hoped that a reasonable and tolerant spirit would prevail within the Catholic Church when science is concerned.

However, over ninety percent of his writing before the *Meditations* was about optic glasses, volcanoes, mountains, rain and snow, about sneezing, bleeding and pumping, about suns, planets and mathematics, and a bit about music. And all those issues were treated in a way that was directly opposed to the Aristotelian approach endorsed by most of the clergy of the time, both catholic and protestant.

This contrast quite spontaneously invited the romantic question: was Descartes sincere when he announced that the goal of the *Meditations* was to defeat the infidels once for all? Wasn't he himself among the greatest living infidels?

This suspicion was incompatible with what I have read as a student in Yugoslavian Marxist literature on Descartes, which seemed to treat his metaphysical investigations at face value, as a deplorable lapse of a great mind into backward obscurity, as a *faiblesse* that could be forgiven because of his otherwise immense contributions, but should not be forgotten and should serve as a warning against the weakening of the real, scientific faith.

However, it was exactly the kind of suspicion that could attract somebody who grew up during the cold war and its fierce ideological conflicts that required strong discipline, especially on the east side of the curtain. Strategies intellectuals used during these complicated times in communist Europe in order to hide what they were really thinking were brilliantly described by the Polish poet and essayist Czesław Miłosz, who named them *ketman* after similar strategies Arthur de Gobineau had discovered in Muslim countries (*Religions et Philosophies dans l'Asie Centrale*, 1866).¹ There is a variety of such strategies, giving rise to a variety of individual moods and mentalities. However, they all have two things in common: the value attributed to one's own opinion, and the decision to preserve it by keeping it secret. Also, there are two ways to preserve the secrecy of one's intimate opinion when it contradicts the official ideology: to be silent about it, and to say the opposite — to profess what the authorities expect one to profess — hoping that only similarly minded people will be able to read between the lines and understand what one really wanted to say. As I understand him, Descartes did both things.

¹ Czesław Miłosz, *Zniewolony umysł* (*The Captive Mind*), Kultura, Paris 1953, chapter III.

The question of the real adversaries

If Descartes wasn't a typical representative of catholic Christianity of the time and lived at a safe distance from catholic authorities, why did he want to defend the basic tenets of their faith? The surprising thing is that the *Meditations* do not seem to fight the infidels at all! In fact, Descartes had mentioned them only twice, in the declarative dedication and in the replies to one of Mersenne's objections who referred to them at first place.² On a closer reading, instead of reasons for suspecting the existence of God, he seemed to be refuting the reasons for suspecting the existence of the external world, and using the proofs of God's existence in order to silence the philosophical sceptic, not atheists or Muslims. Who could that sceptic be, and what was his relation to religion?

In Descartes' time, a philosophical sceptic could have been both a believer and a disbeliever. Fideists held that certainty ultimately eludes rational methods and can be found only through faith, while libertines were suspected of theoretical scepticism and atheism, as well as of improper sexual behaviour. In the *Meditations*, however, Descartes did not mention either of them, and only one of the objectors spoke of libertines.³

The question of the historical identity of Descartes adversaries becomes even more intriguing when one realizes that he never referred to sceptics in *meditations* themselves and mentioned them only once in his replies, where he spoke of them with contempt:

„Now the best way of achieving a firm knowledge of reality is first to accustom ourselves to doubting all things, especially corporeal things. Although I had seen many ancient writings by the Academics and Sceptics on this subject, and was reluctant to rehear and serve this pre-cooked material, I could not avoid devoting one whole Meditation to it.“⁴

According to this explanation, the goal of the *Meditations* is to secure the grounds for sound knowledge, but this goal is not equal to the refutation of classical philosophical scepticism because its main schools are not taken seriously from the very beginning. Doubt is not an impediment to knowledge. On the contrary, it is very useful when it points at

² AT IX-1 2 and 191. All references will be to the classical editions of the french (AT IX-1) and the latin (AT VII) versions published by Adam and Tannery. English translations are taken from Cothingam's english edition of the latin version.

³ It was Arnaud, at the very end of his objections. AT IX-1 288.

⁴ John Cothingam, ed. and transl., *The Philosophical Writings of Descartes* vol. 2, Cambridge, Cambridge University press, p. 94. AT VII 130, AT IX-1 169.

errors where one never expected to find them.⁵ However, such doubt should be greater than sceptic philosophers have ever thought. Their optical illusions and realistic dreams are easy arguments whose refutation does not eradicate all possibilities to err. What is needed is a radical challenge to knowledge whose elimination would finally bring the certainty – what is needed is an *evil genius*. But since this genius was Descartes' own invention, and since he distanced himself from traditional philosophical scepticism, the historical purpose of the *Meditations* could not have been the refutation of this particular school of thought. Was there any real adversary at all, or was it all just one of Descartes many inventions and literary strategies? How could such a question be answered?

In the absence of abundant and coherent historical evidence, the only thing one can do is to follow the arguments themselves and to see where they lead. As it turns out, they lead to an unexpected but familiar door. Here is a very brief summary.

The exposition of the sceptical critique of perception shows that the senses are not a sound source of knowledge.⁶ The subsequent refutation of a radicalized sceptical reasoning with the help of the ontological proof of God's existence shows the superiority of rational thinking. Namely, it is through rational intuitions that we discover that we exist as thinking substances, that there is an unchangeable extended substance in material bodies, and that the two are irreducible to each other.⁷

However, further explanations reveal that the senses are not really unreliable, but insufficient as sources of knowledge, which can be extracted out of them by proper rational analyses (mathematical, physical and metaphysical). Such analysis reveals the true size of distant objects and shows that heat, for example, is not a manifestation of an inner propensity to be hot, but of a faster flow of subtle matter through the cavities of the burning body.⁸

Somehow abruptly, Descartes concludes that all errors about objects of experience are made solely because we lack the time necessary for their careful consideration, and we take them as true images of the way things really are⁹: we believe that the Sun is not bigger than a ball; that the snow is white and the night is black; we believe that the fire is hot and

⁵ See also *Discours de la méthode*, 3, AT VI 28-39.

⁶ AT VII 21-2, AT IX-1 15-16.

⁷ AT VII 30-31, AT IX-1 31-2.

⁸ AT VII 104-6, AT IX-102-3. The precise explanation of heat can be found in chapter three of *Le Monde*.

⁹ AT VII 116, AT IX-1 114.

that sugar is sweet; and we also believe that space in which we perceive no objects is empty.¹⁰

Descartes' choice of examples is highly symptomatic. Namely, if we would strip us of those naive beliefs, we would not only surpass childish opinions about the actual size and shape of distant objects, but would also have to abandon two ontological and epistemological positions held by atomists and Aristotelians.

First of all, the absence of sense data is not the proof of the existence of the void because matter can be transparent,¹¹ which means that we should reject the idea that the world is composed of the void and of atoms. Secondly, colours, tastes and other sensual qualities are not the properties of objects themselves, as the Aristotelians teach, but of mind's reactions to the pressure exerted by external objects on the parts of our body which can be entirely described in terms of quantity.¹²

Finally, the separation of the mind from the body was in itself the most radical of Descartes polemical arguments since it had literally kill the Aristotelian soul by distributing its functions between the two. It was this move that banned various vegetative and organic souls permeating Aristotelian matter, and laid the foundations for a non animistic physics entirely based on analytical geometry.

The real goal

If we follow the arguments, then, the final purpose Descartes wanted to achieve by publishing his *Meditations* was to defeat rival epistemological and ontological projects — the Aristotelian and the atomist one, competing for the same place under the ideological

¹⁰ "There are, however, many other things which I may appear to have been taught by nature, but which in reality I acquired not from nature, but from a habit of making ill-considered judgements; and it is therefore quite possible that these are false. Cases in point are the belief that any space in which nothing is occurring to stimulate my senses must be empty; or that the heat in a body is something exactly resembling the idea of heat which is in me; or that when a body is white or green, the same whiteness or greenness which I perceive through my senses is present in the body; or that in a body which is bitter or sweet there is the selfsame taste which I experience, and so on; or, finally, that stars and towers and other distant bodies have the same size and shape which they present to my senses, and other examples of this kind." Cottingham, p. 56-7, AT VII 1043-4, AT IX-1 101-2.

¹¹ More detailed arguments against the existence of the void can be found in chapter four of *Le Monde*. A good analysis of Descartes longterm struggle against the idea of a void in nature is given in Daniel Graber, *Descartes' Metaphysical Physics*, Chicago, University of Chicago Press, 1992, pp. 136-143.

¹² As Descartes research in *dioptrics* shows, vision itself does not arise when small images that emanate from an object reach the eye, as both atomists and Aristotelians have thought, but through the light which is nothing but a pressure on the eye exerted by the stream of small, shapeless particles, refracted by the surface of the object of vision. See *Dioptrique*, in *Discours de la méthode&Essais*.

umbrella provided by the church, and to show that his own project – a kind of Pythagorean mechanicism – is both theologically and scientifically superior to them.

This reading is corroborated by a well known historical fact. It was in 1633, with a delay of six months, that Descartes heard that the Inquisition has condemned Galileo to silence and home imprisonment for supporting Copernicus' heliocentric idea. At the time, Descartes was finishing his cosmological treatise, *The World*, in which he formulated a specific interpretation of the heliocentric model. As soon as he heard the bad news, he hid his manuscript and informed his colleagues that its publication will be delayed. Instead of *The World*, which would become known only after his death, he published the *Discourse on the method* with the accompanying essays, omitting all cosmological issues, and started to prepare a much more careful and conservative text in which he will develop the metaphysical part of the *Discourse* in a way that should protect his scientific work from all suspicion of heresy. It was the text of the *Metaphysical Meditations*.

Questions of faith

If the real purpose of the *Meditations* was not to defend religion, does that mean that Descartes was not religious? In order to answer this last question more precisely, I will divide it into questions about feelings and questions about belonging.

In his writings and correspondence, Descartes claimed to believe in God, but did not express any of the typical religious feelings, at least not in a typical way. For an east European who grew up during the Cold War, it is a classical case of *ketman* — a case of the intellectual professing to believe in something but omitting to give signs usually associated with this belief. Such attitude, however, can have two different motives: either the intellectual does not believe at all, or he believes in a different way, that is in a different God.

If this reading is justified, Descartes was either lying, or did not tell the whole truth. After more than three centuries, every attempt to resolve this dilemma is inevitably very hazardous. This being so, I will place my bet on Descartes words and suppose that, if he said that he believed in the existence of God, we should as well believe in him.

This, however, invites a further question: were Descartes' feelings Christian? The answer should probably be negative since nowhere in his work can one find any proof that these feelings were attached to objects of Christian faith — the unity of the trinity,

pregnancy without sex, dead man rising from their graves and so on. What we do find is an attempt at a materialistic explanation of the holy secret of the Eucharist (the transubstantiation of bread and wine supposedly occurring during communion) that was meant to show that Cartesian physics is in accordance with the most intriguing tenants of faith.¹³ From the point of the view of the church, however, it was a rather disturbing analysis, and it might have been the reason for which the *Meditations* were put on the Index of forbidden books in Rome twenty two years after their publication (1763), together with other Descartes' books.

It thus seems that Descartes' cautious attitude towards the Catholic church was justified. However, it would be an exaggeration to say that he was not a Catholic. For he grew up in a catholic country, went to a catholic school and recognized the political and social authority of the Catholic church. His religious feelings, whatever they exactly where, did develop in this particular context and could not be entirely detached from it. In a sense, he was a catholic but not a Christian. Less paradoxically, he was a heretic.

¹³ AT VII 354, AT IX-1 336-9.

The Evil Genius – A History of an Argument*

I Introduction

The evil genius argument is the most sophisticated argument used by the philosophical sceptic to question our ordinary knowledge. According to this argument, it is possible that there exists an evil genius who is able to interfere with my knowledge by causing me to have false or inaccurate perceptions which I usually consider true in ordinary circumstances. For example, I observe that I am using a laptop at this moment, therefore I might state „I know that I am using a laptop right now“. In order to support my claim, I can say that the room in which I am touching my laptop is properly illuminated, that there is another person in the room who can confirm that I am using a laptop and so on. Thus, most of us would agree that at this point I can rightly claim that „I know that I am using a laptop“.

However, the sceptic will attempt to question my claim by saying the following argument. It might be the case that there is an evil genius who is deceiving me by imposing me a false impression that I am using a laptop at this moment, while actually that the laptop does not exist, or perhaps I am not even sitting in this room but in a hospital in a deep coma. If this scenario holds, my knowledge claim is false. If it might be the case that I am being deceived by the evil genius at this moment, it might be the case as well that the same genius is deceiving me constantly. If this is correct, it is not possible to assert knowledge claims of any kind at any occasion.

However, many philosophers and non-philosophers would agree that such conclusion is not acceptable, since we have a strong impression that we may be justified when we assert at least ordinary knowledge claims like „I know that I have two hands“. Therefore, if I wish to defend my knowledge claim, I need to show somehow that either

* This paper was supported by the project “Dynamic Systems in nature and society: philosophical and empirical aspects” (evidence 179041) financed by the Ministry of Education, science and technological development of the Republic of Serbia.

sceptic's scenario does not hold or his suggestion regarding the suspension of my knowledge claims is unjustified. Nevertheless, the question is: How could I accomplish such a task? How could I show that I am not being deceived by an evil genius? What kind of evidence I could appeal to? It seems that whichever evidence I employ, the sceptic can always reply – your new evidence is nothing else but another false perception produced by the evil genius.

There were many interesting attempts to reject this sceptical challenge during the last five centuries, and it appears that these debates are still alive. Considering the structure of the responses to the argument and its various formulations, as well as historical and cultural contexts in which these occurred, they could be classified into two sorts – Descartes' approach and approaches offered by contemporary philosophers.

II Descartes

The first philosopher to consider this argument was Rene Descartes. In his *Meditations on First Philosophy* he wanted to investigate if there were any solid grounds on which we could base our system of knowledge.¹ He begins by noticing that it is evident that we know many various things from scientific and non-scientific domains. He was aware that in most cases such knowledge is acquired through senses, calculations, textbooks etc., which are sometimes not completely reliable as sources. It is always possible that I did not perceive something correctly, that I made a mistake during my calculations, or that my textbook is mistaken for whatever reason. This is one of the main reasons why he proposed that we should doubt each and every type of knowledge until we discover a certain base which will grant us that we are not mistaken considering our knowledge claims.²

At this point, the sceptic can argue in two different ways. He might try to show me that my knowledge claims are not well grounded either (1) by appealing to a moderate sceptical scenario which is not compatible with my knowledge claims, or (2) by appealing to a radical sceptical scenario which is also not compatible with my knowledge claims, but which is very difficult or impossible to neutralize. The first route is usually not

¹ Rene Descartes, „Meditations on the First Philosophy: First Meditation“, in: *Key Philosophical Writings*, Hertfordshire, Wordsworth, 1997, pp. 135-139. Translated by Elizabeth S. Haldane and G. R. T. Ross.

² *Ibid.*

troubling us very much since in ordinary circumstances we are able to eliminate such scenarios. For example, if I say „I know that I own an Acer laptop“, the sceptic might claim – „It is possible that the store where you bought your laptop sold out 'Acer' laptops, and then decided to disguise 'Dell' laptop as 'Acer' in order to sell you one“. If this is true, my original knowledge claim would be false. However, it is not quite difficult in this situation to check if the sceptic is wrong. I could contact the manufacturer and check the serial number, for example. Even if it turns out that his objection is true, I will not be upset too much, because the fact that my original knowledge claim was false does not affect in any way my other knowledge claims. The fact that my knowledge claim regarding my laptop brand was false does not suggest that other knowledge claims like „I know that I have two hands“ are false too.

If the sceptic wants to question my knowledge, it would not be quite fruitful to follow the first strategy. Such strategy would probably be both an endless task and I would also be able to defeat him in many cases. This is the main reason why he has to give up such an approach and to appeal to more radical scenarios, like the possibility of the evil genius or the malicious demon. As I have already mentioned, if such scenarios hold, apparently all of my knowledge claims are false.

I will turn now my attention to the solution of this problem Descartes has offered in his *Meditations*. He begins by explaining that, in a certain way, our knowledge is based on ideas which we perceive in our minds. For example, I am perceiving now that I am sitting at my desk. There are three types of ideas: (1) The ones I create by myself by using my imagination, (2) ideas produced by my perception and (3) ideas which are in a certain way innate.³ The main feature of (1) is that they are not certain and reliable, because I can create an idea which does not correspond to any object in the world (the idea of the unicorn). In the same sense, (2) can sometimes be true, depending on various conditions like reliability of my senses in the given circumstances (it is possible that my senses were not accurate when I saw an airplane in the sky, because it might turn out that the object I saw was actually a bird). The chief feature of the last type (3) is that they cannot be false, i.e. they are clear and distinct which we can not say about (1) and (2). One of the examples is Descartes' famous *cogito*. Even if it were the case that I am mistaken about many things,

³ *Ibid.*, p. 150.

I could not be mistaken about myself who registers such mistakes, for there must be a subject (*I*) to register mistakes in order to say „I am mistaken“.⁴

Now, it is somewhat easier for Descartes to attempt to eliminate the possibility of the evil genius. Even if it were the case that I am constantly being deceived by such creature, it only follows that there must exist a subject who is being deceived, and that subject is me. As I already said, I could be wrong about the existence of many things, but not about the existence of myself or, to put it in other words, I could not be wrong about the existence of my mind. As long as I think and realize this fact, it is certain that at least *I* exist.

Apparently, Descartes has fulfilled his task by showing that there is certain kind of knowledge immune to the evil genius argument. In fact, the argument could only help us rather than disturb us. For whenever the sceptic attempts to challenge my knowledge claims by presenting me the possibility of the evil genius who might be deceiving me, I need to doubt in my knowledge claims. But, when I do that, I realize pretty fast that it is certain that there must *exist* a subject who is being deceived, i.e. myself.

It would be interesting to briefly mention some of the many reactions to Descartes' theory by philosophers and theologians who lived at his time. For the most part, Descartes' thought was not very well accepted, and eventually it was banned by the Church. Some contemporary interpreters (Ted Schmaltz) believe that the Church had several reasons for such action.⁵ According to Schmaltz, what Church did not like about Descartes' *Meditations* was his interest in reviving the scepticism combined with postulating an independent researcher who is always pursuing truth relying mostly on his own critical approach to philosophical issues and independently of the (religious) dogmas. I believe that there might be another further similar reason for the prohibition. If we carefully read Descartes' works, we shall notice there is a certain switch from the Church and its dogmas as the main source of our knowledge, to the independent researcher and his own cognitive faculties. Not only that the researcher is independent of dogmas, but he also possesses a completely new methodology of research, according to which one needs to start by universally doubting all of her knowledge, and then to arrive at certain knowledge all by herself, without any influences of external authorities.

⁴ *Ibid.*, pp. 141-142.

⁵ Ted Schmaltz, „Seventeenth-century Responses to *Meditations*“, in: Stephen Gaukroger, *The Blackwell Guide to Descartes' Meditations*, Oxford, Blackwell Publishing, 2006, pp. 193-203.

III The difference between Descartes' and contemporary cultural and historical contexts

One of the main reasons why Descartes decided to formulate his argument of the possibility of the evil genius is because the malicious genius or daemon is directly opposed to God who is characterized as good and benevolent. As we know, most of the philosophers of his time were usually trying to relate their works to theology or theological assumptions. Descartes fulfilled this task by introducing a malicious being who is constantly deceiving him, and also tried to show why there is a God who is doing the opposite.

The modern form of the evil genius argument is based on a completely different cultural and historical context. Nowadays, Church and religion do not affect philosophical investigations as much as they did during Descartes' life. That is one of the main reasons why contemporary philosophers are not „obliged“ to use arguments related to theological subjects, but instead look for their examples in various domains like pop-culture, science fiction in movies and books, various suspicious scientific theories and so on. Apart from that, modern philosophers can use many important discoveries in epistemology, philosophy of language and logic which can help them to formulate and investigate the evil genius argument more accurately. As I will mention soon, some contemporary epistemologists treat this argument as a mere (intentional) misuse of our ordinary language.

In the most common contemporary form of the argument, the evil genius is depicted as an evil neuroscientist who keeps my brain in a vat, attached to a supercomputer which is capable of producing all the perceptions and thoughts I usually have. The scientist is therefore able to cause me to believe that I am now sitting at my desk, while I actually do not have body parts except for my brain which is being kept in a vat.

IV Contemporary responses to the evil genius argument

As I already said, there are many interesting strategies found in modern literature to solve the sceptical challenge. I will try to present some of them which I find the most interesting.

One of the responses was offered by the so called theory of relevant alternatives. According to this account, when the sceptic presents me a scenario which aims to question my knowledge about some fact, he will basically employ the same strategy as Descartes' malicious genius does. Such scenario represents an alternative which, if it turns out to be correct, I need to suspend my knowledge claim. Let us suppose that I possess some knowledge about paintings and that I visit an exhibition of Picasso's works. When I stand in front of one of those paintings, I might say „I know that this is Picasso's painting“. If I am asked to support my claim, I can say for example that I learned earlier that this particular painting is painted by this artist, that I can recognize his peculiar style of painting, that there is a tag under the painting with Picasso's name and so on. If the sceptic visits the exhibition too, he could challenge my claim by saying „what you are seeing right now might be actually nothing else but a very well made copy of his painting. The genuine painting might have been stolen last night, and the gallery decided to display this copy in order to avoid public embarrassment.“ If this alternative holds, my knowledge claim is of course false: I do not know that this is Picasso's painting. Thus, if I wish to keep my claim I need to eliminate this alternative. In this particular case, as we saw earlier when I was discussing Descartes, this should not be a tough challenge. For instance, I could ask my friend who is the director of the gallery who wouldn't lie to me if this particular work is genuine or just a copy. If the director confirms that I am looking at the genuine painting, the sceptical alternative is eliminated and the sceptic is defeated. If not, I am the one who is defeated.

However, things get much more complicated if the sceptic decides to employ a radical alternative rather than a moderate one which I just described, because these are impossible to eliminate. He might claim that it is possible that right now the evil scientist is using his supercomputer to stimulate my brain so that I have an impression that I am looking at a painting, while both me and that painting do not exist. Does it mean that I must accept the fact that since I am never able to eliminate this alternative I can never be sure about anything? According to the theory of relevant alternatives and one of its main

proponents – Fred Dretske – this would be a wrong conclusion.⁶ This theory, roughly said, maintains that my duty is only to eliminate relevant alternatives and that I should pay no attention to irrelevant ones. In the previous example, the possibility that I am looking at the copy of the painting might be considered as a relevant alternative. In that case, I am obliged to show that this scenario is false. However, the possibility that I am being deceived by the evil genius regarding the painting constitutes an irrelevant alternative, for such scenario is highly improbable. Therefore, I am not obliged to eliminate this alternative.

Now we can ask – what makes an alternative relevant? Dretske's theory suggests that an alternative is relevant if there exists a reasonable probability that it is true or might become true. For instance, if I am sitting in a dark room and say „I know that this piece of paper is white“, the sceptic might object that the piece of paper is perhaps yellow. I must admit that there is some reasonable probability that this alternative is correct which will make it then relevant, because the room in which we are sitting is not well illuminated. Once again, I can prove the sceptic to be wrong by turning the lights on and showing him that the piece of paper is really white and not of some other colour.

What is the then status of the evil genius argument and of similar radical alternatives? According to this account, the probability that at this moment I am being deceived by a malicious genius or evil scientist is very low, for such probabilities rather belong to the domain of fiction or mere probabilities than to the domain of probability. If this is correct, then the evil genius argument belongs to the category of irrelevant alternatives. Since this is so, I do not have an obligation to eliminate it, I do not even have an obligation to consider it seriously as a possible threat to my knowledge. Therefore, I should not be concerned with such scenarios and my knowledge is safe.

This theory gained much popularity probably because it seems to us at the first glance as quite simple and effective. However, we could ask now – how do we determine or measure the probability that an alternative might turn out to be true? How many times an alternative must be realized before we declare it ir/relevant? Is it sufficient to encounter a sceptical scenario only once in order to say that such particular scenario has reasonable probability to hold, or should we encounter them more often? Also, we can ask if the

⁶ Fred Dretske, „Pragmatic Dimension of Knowledge“, *Philosophical Studies*, vol. 40. No. 3, 1981, pp. 363-378.

probability varies from context to context? For example, does the same alternative have an equal level of probability in one gallery but not in the other?

This is the main reason why another group of philosophers called *contextualists* tried to approach this problem in a different way. According to them, whenever we make assessments about our knowledge claims, we do that in relation to the context of the utterance of knowledge claims. Roughly said, contexts include certain epistemic standards upon which we base our assessments.⁷ For example, when I state the following claim „I know that I am looking now at a painting painted by Picasso“, I must determine what is the context of my utterance, that is, wheather this context includes low or high epistemic standards. If the former is the case, if my claim is made in the context with lower epistemic standards which are relatively easy to satisfy, then my claim is true. On the contrary, if the latter is the case, i.e. the context includes higher epistemic standards which are much more difficult or in some cases impossible to satisfy, my claim might be false.

Does it mean then that one sentence can be true and false at the same time, but in different contexts and relative to different epistemic standards? Roughly said – yes. In the context which includes lower epistemic standards my claim „I know that this is Picasso’s painting“ will be true, while on the other hand the same claim will be false in the context which includes higher epistemic standards. Following the same pattern, when I say „I know that I am not being deceived by the evil genius“ in the context with lower epistemic standards my claim will be true, but then again false in the opposite context.

At this point, we could say that the sceptic actually misunderstands or abuses contexts – he is interpreting our knowledge claims which we make in contexts with lower epistemic standards in the light of higher standards. What the sceptic must not do is to combine higher with lower contexts as it suits him. He is also not permitted to transfer conclusions from one context (for example, my conclusion made in higher context in which I claim that I do not know that I am not being deceived by the evil genius) to support his sceptical claim made in the lower context. This is not allowed because epistemic standards from one context do not apply to another.⁸

Contextualists proceed by attempting to explain with more details why does the sceptic mistakenly manage to switch different contexts. In order to complete this task, as well as to strengthen their positions further, contextualists often appeal to basic ideas of

⁷ For example, see Tim Black, „Contextualism in Epistemology“, available online at <http://www.iep.utm.edu/contextu/> (last accessed 01/01/2014).

⁸ *Ibid.*

the theory of relevant alternatives. One of them, David Lewis, tries to explain the switch by formulating the so called *rule of attention*.⁹ According to this rule, an alternative becomes relevant as soon as one of the participants in a conversation mentions it, or as soon as a person becomes aware of it. The alternative that I might be looking at a copy of Picasso's painting instead of the original painting becomes relevant when the sceptic mentions such a possibility. In this case, we have automatically shifted to a higher context, i.e. to a context which includes higher epistemic standards. As in earlier cases, such alternative is not particularly troubling because I can easily eliminate it. In the case of the evil genius, however, such elimination is not possible. Lewis suggests that since we wish to preserve our knowledge, we should simply never mention any such radical alternatives. According to this account, we are aware that evil genius might be deceiving us, but we are also aware that we are not able to prove this possibility wrong, so we should just remain silent.¹⁰

I believe that Lewis' suggestion is not sound, for it is obvious that the sceptic will not remain silent and that we have an impression that he is not defeated by the strategy proposed by Lewis. He will attempt to attack our knowledge, and once he does that, we can hardly remain silent.

Another explanation of the sceptic's misuse of contexts has been offered by Gail Stine.¹¹ In order to show her theory, it might be useful to present the sceptical challenge in the form of the so called *sceptical paradox*. Consider the following claims:

- (1) I know that I have two hands.
- (2) If I know that I have two hands, then I know that am I not being deceived by an evil genius.
- (3) I do not know that I am not being deceived by an evil genius.
- (4) I do not know that I have two hands.

It appears that each one of these claims holds when we consider them separately. However, it also seems that they are not consistent when we consider them together. If we take a closer look at these claims, we can see that they do not include the same context with the same set of epistemic standards. Let us first consider (1). As we already saw, such claim is usually true in contexts with lower epistemic standards which make this

⁹ David Lewis, „Elusive Knowledge“, *Australasian Journal of Philosophy*, 74, 1996, pp. 549-567.

¹⁰ *Ibid.*

¹¹ Gail Stine, „Scepticism, Relevant Alternatives and Deductive Closure“, *Philosophical Studies*, 29, 1976, pp. 249-261.

claim true. On the other hand, we can also see that the context has been switched in (2) so that we are now employing context which includes higher epistemic standards. Of course, in that case, he will proceed to assess the claim (3) in the same light, by using higher epistemic standards, which will be true by such standards. According to Stine, this is actually an instance of equivocation. Sceptic is not allowed to switch contexts in these claims, he is obliged to use just one of them. But the sceptic can not fulfil this requirement. If he only insists on employing the context with lower epistemic standards in each of these claims, then (1) and (2) will be true but (3) will be false, and therefore my knowledge of ordinary facts will remain untouched. On the other hand, if he decides to employ a context with higher epistemic standards, then (1) will be false, while (3) will be true. In that case, the sceptical suggestion that I do not know that I have two hands will hold. However, according to some philosophers, if the contextualist interpretation of the sceptical challenge is correct, the sceptical suggestion should not worry us, because we never intend to claim that we assess ordinary knowledge claims by high epistemic standards. When I say in ordinary circumstances „I know that I have two hands“, according to some contextualists, I rely on ordinary or lower context which includes lower epistemic standards which are easier to satisfy. Therefore, if the sceptic intends to question my knowledge of ordinary facts by appealing to radical sceptical hypotheses, he needs to show somehow that I am not able to exclude the evil genius possibility in ordinary contexts with lower epistemic standards. However, as we already saw, if we accept contextualism, such task would be very difficult to complete.

Conclusion

We have seen that the evil genius argument is a very striking argument which tends to question all of our knowledge. Descartes has neutralized it by appealing to some remarks from his philosophy of religion and philosophy of mind. I believe that contemporary philosophers are in a better position to formulate and cope with this argument. Contextualists believe they have succeeded in showing in a more or less convincing way how can we prove the sceptic to be wrong. However, there are many other contemporary philosophers who are not happy with their solution, charging them that they relativize knowledge too much. The other charge, which I find quite plausible, states that contextualists are not able to explain the so called asymmetry of contexts: if we accept

contextualism, once we leave the debate with the sceptic, we should be able to simply forget about all the issues which we discussed with him. However, this seems very unlikely: it does not sound quite plausible that I am not bothered at all by the evil genius possibility when I „return“ to a lower context.

I believe that debates about this argument are still pretty much alive. Unlike Descartes, contemporary philosophers are equipped with important discoveries mainly in epistemology and philosophy of language, which could help them to indicate some further misuses or fallacies committed by the sceptic or the philosophers who attempt to defend our ordinary knowledge claims. Perhaps it will turn out that the issue with the possibility of a malevolent deceiver and the requirements for certain knowledge is nothing else but a linguistic misunderstanding between us and the sceptic.

References

- Black, Tim, „Contextualism in Epistemology“, available online at <http://www.iep.utm.edu/contextu/> (last accessed 01/01/2014).
- Descartes, Rene (1997) „Meditations on the First Philosophy: First Meditation“, in: *Key Philosophical Writings*, Wordsworth, pp. 135-139. Translated by: Elizabeth S. Haldane and G. R. T. Ross.
- Dretske, Fred (1981) „Pragmatic Dimension of Knowledge“, *Philosophical Studies*, vol. 40, n° 3, pp. 363-378.
- Lewis, David (1996) „Elusive Knowledge“, *Australasian Journal of Philosophy*, vol. 74, pp. 549-567.
- Schmaltz, Ted (2006) „Seventeenth-century Responses to *Meditations*“, in: Stephen Gaukroger, *The Blackwell Guide to Descartes' Meditations*, Blackwell Publishing, pp. 193-203.
- Stine, Gail (1976) „Scepticism, Relevant Alternatives and Deductive Closure“, *Philosophical Studies*, 29, pp. 249-261.

Cartesian Scepticism*

My aim in this paper is to show how the Cartesian sceptical strategy, developed in Descartes' First Meditation, may be generalized so as to result in some different and, in my opinion, more profound form of philosophical scepticism.

Generally speaking, philosophical scepticism has to do with the necessary conditions for knowledge. For any C , conceived as a necessary condition for knowledge, the sceptic argues that we are not able to satisfy C and concludes that, for this reason, we cannot claim to have knowledge.

I will stick to the familiar definition of knowledge, which lists its three necessary (jointly sufficient) conditions: truth, belief, and justification. The Cartesian sceptical strategy is directed only at the truth condition, since it appeals to some possibilities of mistake that cannot apparently be ruled out and thus may undermine (almost) all of our knowledge claims. I will try to show that the well known Cartesian hypotheses, which invoke such possibilities, may be interpreted or subtly revised so that they can be used to attack the other two conditions listed in the traditional definition, thus providing a ground for two non-Cartesian forms of scepticism.

1. Cartesian scepticism: the quest for certainty and the closure principle

The main target of philosophical scepticism is the common sense view that we know many things: I know that I have hands, I know that I am living in Belgrade, we know that we are human beings, and so on.¹ However, as soon as we engage in a systematic

* This paper was supported by the project "Dynamic Systems in nature and society: philosophical and empirical aspects" (evidence #179041) financed by the Ministry of Education, science and technological development of the Republic of Serbia.

¹ Surely, this view prevails among philosophers as well, as it is shown in the survey recently conducted by Bourget and Chalmers (Bourget and Chalmers, 2013). They surveyed many professional philosophers in order to find out their views on thirty central philosophical issues. One of them was the question of choice between idealism, scepticism, and non-sceptical realism, with the following result:

philosophical investigation of knowledge, in the way that Descartes did, for instance, we face a compelling argument that we know almost nothing.

Cartesian sceptical arguments start with the so-called infallibilist thesis that knowledge requires absolute certainty, i.e., that it must be, by definition, infallible. At first glance, the thesis seems plausible: if someone claims that she knows some proposition p , and yet admits that she cannot eliminate a certain possibility q in which not- p , it certainly seems as if she has granted that she does not know p after all. As Peter Unger, David Lewis, and some other philosophers have pointed out, to speak of fallible knowledge, or of knowledge despite uneliminated possibilities of error, seemingly sounds contradictory.²

One possible reading of the infallibilist thesis is that one knows p on the basis of evidence e only if e entails p , or, in other words, only if e is such that there is no q consistent with e and inconsistent with p .

From the common sense point of view, this reading is doomed to a radical sceptical conclusion. In everyday life, especially if we take contingent propositions about the external world into account, this requirement can never (or hardly ever) be met - whenever we claim to know any p , there are almost countless ways in which we could be wrong (conspiracies to deceive, illusions, hallucinations, dreams, and so on).

A simple way to avoid this immediate sceptical result is to reject the infallibilist thesis. This move amounts to accepting a fallibilist theory of knowledge, which holds that one can know p on the basis of e where e only makes p probable. When we claim that we know p , we certainly imply that we are not wrong about the truth of p , or at least it is not obvious that we are implying that we have successfully eliminated *all* possibilities for error. How on earth could any human being rule out *every* possibility of mistake before being in a position to claim that she or he knows something?

Unfortunately, scepticism cannot be so easily discarded. Although this fallibilist move apparently squares with our strong common sense intuition that we know many things, due to other principles that are very difficult to reject it is forced to confront scepticism at least in the form of a persistent paradox. Cartesian scepticism, as developed by Descartes in his *Meditations* (especially in the First Meditation), provides an

Non-sceptical realism 81.6 [Accept: 61.4%, Lean toward: 20.2%]

Other 9.2 [Accept another alternative: 2.6%]

Scepticism 4.8 [Lean toward: 3.0%, Accept: 1.8%]

Idealism 4.3 [Lean toward: 2.7%, Accept: 1.6%].

² Compare Peter Unger, *Ignorance: A Case for Scepticism*, Oxford, Oxford University Press, 1975, and David Lewis, "Elusive Knowledge", *Australasian Journal of Philosophy*, 74, 1996.

illustrative example.

As we know, Descartes granted that knowledge requires certainty, but his reading of the infallibilist thesis was more subtle. He started from the minimal knowledge requirement, consistent both with infallibilism and fallibilism, which states that when we claim to know *p*, we should be able to rule out at least the possibilities of mistake that we encounter. But, instead of resorting to the indisputable fact that, whenever we claim to know *p* on the basis of *e*, there are *some* possibilities that could render us wrong, he focused on *particular* error possibilities such that, once we are faced with any of them, it turns out that we cannot possibly rule them out, and are thereby forced to accept the sceptical conclusion. In other words, according to the Cartesian sceptic, the problem is not that there are always *some* possibilities of mistake; it is that there are *particular* error possibilities that we are not able to exclude.

The main principle on which Descartes based his well known method of doubt, which generates the sceptical argument, is presented in his earlier work *The Discourse on Method* as a methodological rule which, among others, governs epistemic rationality:

“The first was never to accept anything as true if I did not have evident knowledge of its truth: that is, carefully to avoid precipitate conclusions and preconceptions, and to include nothing more in my judgements than what presented itself to my mind so clearly and so distinctly that *I had no occasion to doubt it.*”³

Let’s remind ourselves how Descartes arrived at this rule.

At the very beginning of the First Meditation Descartes confesses that he has realized that he has made many mistakes in the past: “Some years I was struck by the large number of falsehoods that I had accepted as true in my childhood, and by the highly doubtful nature of the whole edifice that I had subsequently based on them” (*Med.*, p. 12: 17).⁴ From this insight he has drawn the moral that “... I should hold back my assent from opinions which are not completely certain and indubitable just as carefully as I do from those which are patently false...” (*Med.*, p. 12: 18). In other words, he realizes that he should be more cautious and reject every belief for which he finds “at least some reason

³ Descartes, 1985, p. 120: 18 (my italics).

⁴ Throughout the paper, “*Med.*” refers to the English translation *Meditations on First Philosophy, With Selections from the Objections and Replies*, translated and edited by John Cottingham, Cambridge, Cambridge University Press, 1996 (f. p. 1986).

for doubt” (*Med.*, p. 12: 18).

Descartes’ methodological rule may be generalized and expressed in the following way:

In spite of all available evidence e , one should refrain from claiming to know p if one finds at least some reason for doubting that p is true.

In light of this rule, we may put forward the Cartesian Knowledge Requirement:

In order to claim to know p , one should be able to rule out every possibility of mistake, i.e., every alternative q inconsistent with p , that one encounters.

This requirement sets out a framework for a dialectical debate between those who claim to know a lot and the Cartesian sceptic. Suppose that, on the base of evidence e , we claim to know p . Relying on the requirement, the sceptic draws our attention to hypotheses which are such that (a) they specify possible situations in which not- p ; (b) by their very nature are consistent with e and cannot be ruled out by any piece of e that we might want to use; and (c) are incompatible with the truth of p , such that, if correct, they prevent us from knowing p . According to this requirement, since such hypotheses invoke possibilities of error that are inconsistent with our knowledge claim, it seems that we have to rule them out in order to have knowledge that p .

We can illustrate this point with the well known Dretske’s zoo example.⁵ Suppose that we are visiting the Belgrade zoo. We are walking around watching different animals, until we stop in front of a striped animal; in daylight, we see it is a zebra. Still, there is the possibility that it is a mule, for some reason cleverly painted so as to look like a zebra. This alternative is consistent with the visual evidence we have in the given circumstances, and it undermines our knowledge that it is a zebra, since the hypothesis that it is actually a cleverly painted mule is incompatible with the assertion that it is a zebra.

Of course, even if we don’t know that the striped animal is a zebra at the moment when we look at it, we are nevertheless able to find out subsequently whether it is a zebra or a cleverly painted mule. On the other hand, the aim of the Cartesian sceptic was to discover particular possibilities of error which we would never be able to exclude and which could provide the ground for universal doubt concerning all (or nearly all) ordinary knowledge claims, especially concerning our knowledge of the external world.

⁵ See Fred I. Dretske, Epistemic Operators, *Journal of Philosophy*, vol. 67, 1970, p. 1016.

Descartes' dreaming hypothesis describes such a possibility. While sitting by the fire and writing his *Meditations*, Descartes remarks:

How often, asleep at night, am I convinced of just such familiar events - that I am here in my dressing-gown, sitting by the fire - when in fact I am lying undressed in bed! ... As if I did not remember other occasions when I have been tricked by exactly similar thoughts while asleep! As I think about this more carefully, I see plainly that there are never any sure signs by means of which being awake can be distinguished from being asleep. The result is that I begin to feel dazed, and this very feeling only reinforces the notion that I may be asleep. (*Med.*, p. 13: 19)

Another possibility is depicted by the so-called evil demon hypothesis. Allowing that God, as a perfect being, is by definition so good that He would not wish us to be systematically deceived, Descartes supposes that "... not God, who is supremely good and the source of truth, but rather some malicious demon of the utmost power and cunning has employed all his energies in order to deceive me. I shall think that the sky, the air, the earth, colours, shapes, sounds and all external things are merely the delusions of dreams which he has devised to ensnare my judgment. I shall consider myself as not having hands or eyes, or flesh, or blood or senses, but as falsely believing that I have all these things." (*Med.*, p. 15: 22–23)

Let us put Descartes' remarks in more general terms. While forming beliefs about the external world, we can only appeal to our perceptual experience. We believe that, in normal circumstances, ordinary physical things around us cause our perceptions, and through these we acquire knowledge of the external world. Nevertheless, it seems that all our perceptions may be equally well explained by the dreaming hypothesis, or by the evil demon hypothesis, and since nothing in our subjective experience can reveal to us that we are not simply dreaming delusive dreams, or not being deceived by a powerful malicious demons, there is the possibility that the world is radically different from the way it appears. Thus, the Cartesian sceptic concludes that we lack knowledge about the external world, since we cannot exclude the possibility that, according to the dreaming or the evil demon hypothesis, the supposed things around us are non-existent or at least entirely different from the way we take them to be.

Here we need a few clarifications. First, the sceptic does not need to argue that her hypothesis is true. Moreover, Descartes permits that she doesn't have to provide any further reason why she thinks it might be true in any given circumstances. A mere

possibility of the truth of sceptical hypotheses serves the sceptic's aim quite well.⁶ Yet the sceptical hypothesis purports to provide an alternative causal explanation of our beliefs by supposing that their causal origin and history are radically different from the way we take them to be. At the same time, it is the hypothesis that, if true, explains how it happens that our beliefs are false in given circumstances.

Second, by pointing to the particular possibility of mistake, the sceptic intends to dispute the assertability, rather than the truth of our knowledge claim. In other words, the sceptic does not want to assert that we don't know *p*, but only that, according to the Cartesian knowledge requirement, the burden of proof is on our shoulders and, unless we are able to show that her hypothesis is false, we can't rightly claim to know *p*. In this way, we are involved in a dialectical debate the outcome of which depends on our prospects in proving that this error possibility does not actually obtain.

Finally, and most importantly, the Cartesian strategy relies on the epistemic version of a logical principle that we cannot so easily dismiss. This is the so-called *closure principle*, which states that knowledge (and possibly justification) is closed under entailment:

If one knows *p* and one knows that *p* entails *q*, then one may come to know *q* simply by competently inferring it from *p*.⁷

While the closure principle strikes most philosophers as compelling, it presents difficulties for the common sense fallibilist view. As fallibilists, who reject the theses that knowledge must be infallible, we can allow that we may know *p* on the basis of *e*, even though *there is a* proposition *q*, consistent with *e* but inconsistent with *p*. But if we are faced with *q*, and at the same time know that *p* implies not-*q*, then even as fallibilists we are in trouble of running afoul of the closure principle.

Cartesian hypotheses may be presented in this way. Consider the claim that, now, while I am writing this paper, I know that I am sitting in a chair. The dreaming and the evil demon hypotheses describe alternatives, consistent with the perceptual evidence on which I base my claim, but inconsistent with the proposition "I am sitting in my chair", and it becomes fairly obvious that, in order to know that I am sitting in a chair, I need to

⁶ Curley holds that Descartes made this concession to the sceptic by purpose. E.M. Curley, *Descartes Against the Skeptics*, Oxford, Blackwell, 1978.

⁷ The proviso 'by competently inferring' is needed in the standard formulation of the principle, since it might happen that someone knows *p* and that *p* implies *q* without inferring *q*. See Timothy Williamson, *Knowledge and its Limits*, Oxford, Oxford University Press, p. 117; John Hawthorne, *Knowledge and Lotteries*, Oxford, Clarendon Press, 2004, p. 32.

know that I am not lying asleep in my bed and dreaming, or that I am not being deceived by an evil demon. But since my experience would be just as it is even if I was deceived by a dream or an evil demon, we are reluctant to say that I can possibly know that I am not so deceived. But then I do not know that I am sitting in a chair either.

Relying on the closure principle, this argument leads to the Cartesian sceptical paradox. Where p is a proposition we would ordinarily claim to know, and q is a sceptical hypothesis whose denial is entailed by p , the sceptic employs the claim:

S does not know not- q

in conjunction with:

If S does not know not- q , then S does not know p

to infer:

S does not know p ,

which contradicts the starting claim:

S knows p .

These propositions are independently plausible, but, when combined, are obviously inconsistent and constitute a paradox. Some philosophers (most notably Descartes and Moore) have attempted to resist scepticism by denying the claim that we do not know not- q . Others (for instance Dretske and Nozick) have argued that we should give up the closure principle, allowing that we may know p even if we don't know not- q . And the third, now very popular proposal is contextualist: both the ordinary knowledge claim that we know p , and the sceptical claim that we don't know not- q , are true, but in different contexts and according to different epistemic standards. I will not, however, discuss these here.

2. General Knowledge Requirement

Much recent work on philosophical scepticism treats Cartesian scepticism as relying on the closure principle.⁸ Yet I think that the knowledge-undermining strength of the

⁸ For instance: Dancy, *An Introduction to Contemporary Epistemology*, Oxford, Blackwell, 1985; Stanley Cohen, "Skepticism, Relevance, and Relativity", in: Brian McLaughlin, ed., *Dretske and His Critics*.

Cartesian sceptical hypotheses is more far-reaching. In order to show how far it might reach, I will first point out how Cartesian sceptical strategy can be generalized so as to provide the ground for some non-Cartesian forms of philosophical scepticism, which will turn out to be even more profound than Cartesian scepticism itself.

Let's go back to the concept of knowledge. Any plausible definition of this concept should produce a list of necessary conditions that are jointly sufficient for knowledge. Whatever definition is offered, the sceptic will accomplish her goal if she finds a hypothesis that, were it true, would hinder the fulfillment of at least one of the necessary conditions listed in the definition. Take, for instance, the traditional definition of knowledge according to which a person *S* knows *p* if and only if *p* is true, *S* believes that *p*, and *S* has a *justified* belief that *p*.⁹ The basic Cartesian strategy is to undermine one's knowledge that *p* by pointing out the particular possibility of mistake as an alternative which, being inconsistent with *p*, prevents the satisfaction of the truth condition. However, it seems that Cartesian sceptical alternatives may be adjusted in such a way as to attack the other two conditions. Hence, we may take as an alternative any sceptical hypothesis that is inconsistent with a knowledge claim "*S* knows *p*" and, depending on the necessary condition for knowledge at which they are targeted, divide them into (1) T-alternatives (directed at the truth condition), (2) B-alternatives (directed at the belief condition), and (3) J-alternatives (directed at the justification condition).

In light of this, the Cartesian Knowledge Requirement may be generalized.

General Knowledge Requirement: In order to know *p*, one has to be in a position to rule out at least those possibilities [alternatives] which one recognizes as *incompatible with one's knowing* that *p*.

Both Descartes' sceptical hypotheses might be interpreted in a slightly different way so to undermine knowledge through the belief or the justification condition, rather than through the truth condition. It will turn out that their knowledge-destroying capacity does not directly stem from the closure principle and their supposed inconsistency with the truth condition.

We have already illustrated the way in which T-alternatives work against ordinary knowledge claims. Without going into detail, in the rest of this paper I will outline how

Cambridge, Mass., Basic Blackwell, 1991; K. Falvey and J. Owens, "Externalism, Self-Knowledge, and Skepticism", *The Philosophical Review*, vol. 103, n°1, 1994.

⁹ It should be noted that my goal here is not to defend the traditional analysis; the same line of reasoning that follows would apply whatever definition is offered.

the Cartesian hypothesis can be construed so as to invoke B- and J-alternatives.

3. Cartesian hypotheses as B-alternatives

In light of the general knowledge requirement and traditional definition of knowledge, the second way in which the philosophical sceptic may undermine knowledge is to put forward an alternative that, if actualized, would prevent us from satisfying the belief condition.

Descartes did not consider this option seriously because he started from the assumption that the contents of all of one's thoughts are completely determined by "internal" facts about what is going on inside one's mind (head) and, moreover, that one can never be mistaken in identifying his own mental states (including his beliefs). The first claim is known as the thesis of semantic internalism, while the other is known as the thesis of epistemically privileged access.

It seems that the original Cartesian alternatives essentially depend on semantic internalism. For instance, in the evil demon scenario, the Cartesian sceptic takes for granted that it is possible that, in a world which is by hypothesis completely unlike ours, a person may have the same beliefs as she has in the actual world; the only difference between a person's beliefs in the actual and in an evil demon's world is supposed to be in their truth values, not their contents. In other words, the Cartesian sceptic assumes that even in situations in which the causal origins of almost all our beliefs are totally different, due to some internal facts we could have the very same beliefs that we have in the actual world.

Following Putnam, some philosophers (for instance, Brueckner, Warfield and Forbes) have disputed this assumption, charging the Cartesian sceptic with inconsistency. According to them, if the causes of our beliefs were from the beginning totally different from those in the actual circumstances, at least some of our beliefs about the external world would have to differ in content from the analogous beliefs we actually have. In fact, they reject semantic internalism and adopt the opposite view on mental content, namely semantic externalism.

Semantic externalists hold that the contents of at least some of one's thoughts (for example, thoughts expressed by words which directly refer to physical objects, such as indexicals, proper names, or terms for natural kinds) are not entirely determined by

“internal” facts about what is going on inside one’s mind (head), but are at least partly determined by “external” facts, such as the nature of the objects one has been in causal contact with.

This is an appealing view. Imagine that we have never been in proper causal contact (visually or kinesthetically) with hands, as physical objects. How could we ever form a concept of a hand, or any belief about hands, if we hadn’t been causally connected with hands in the proper way?¹⁰ It seems that a person systematically deceived by an evil demon, thereby not having been in contact in the proper way with real hands, cannot think or believe anything about real hands. When this person believes the proposition expressed by the sentence “Here is a hand”, then, given that we are not deceived by an evil demon, she is not believing the same proposition that those words would express were they in our mouths. Putnam has listed few suggestions as to what might be the reference for a brain-in-a-vat’s when they “use” the word “hand”. The most suitable suggestion for Descartes’ evil demon scenario would be that it refers to hands-as-the-idea in that person’s mind. This suggestion is in the spirit of semantic externalism, because, as Putnam notes, “there is a close causal connection between” the person’s use of the word “hand” in her language, and the idea of the hand created in her mind by the evil demon. Importantly, on this suggestion, the person ends up believing that something is true when she thinks “Here is a hand”, because she is referring to a hand-as-the-idea inside her mind.

Of course, it might be questioned whether a person in the evil demon scenario can have any *true* thoughts at all, but as a response to Cartesian scepticism it would suffice to accept a weaker, negative claim to the effect that, due to differences in external factors, she cannot have a great deal of the thoughts that we actually have. The main point is that she is not thinking the same thoughts as us when we are thinking about this hand, given

¹⁰ Putnam is dealing with the so-called brain-in-a-vat hypothesis, a modern version of Descartes’ evil demon hypothesis. Starting from semantic externalism, he has claimed to have “an argument ... that shows we are not brains in a vat” (Hilary Putnam, *Reason, Truth and History*, Cambridge, Cambridge University Press, 1981, p. 8), and that promises to cut off Cartesian scepticism via a denial of the first premise of the sceptical argument. Putnam himself was largely uninterested to develop this argument as an answer to scepticism, but others have seized on it as an anti-sceptical weapon (Brueckner, Warfield, Forbes, etc). This strategy resembles Descartes’ response to the evil demon argument; as DeRose noticed, externalist semantics has replaced Descartes’ God as a sort of guarantee that the majority of our beliefs about the external world is true (Keith DeRose, “How Can We Know that We’re Not Brains in Vats?”, *The Southern Journal of Philosophy*, vol. 38, n° 51, 2000, p. 143). Here I do not want to deal with this kind of anti-sceptical strategy; rather, I would like to point out how semantic externalism construed as an account of “the mind/world relationship” (Putnam, p. 31) gives rise to a second-order scepticism which Brueckner calls *scepticism about knowledge of content* (A. Brueckner, “Semantic Answers to Skepticism”, in: K. DeRose and T. A. Warfield, eds. *Skepticism, A Contemporary Reader*, New York – Oxford, Oxford University Press, 1999, pp. 48-9; see also Brueckner, “Scepticism about Knowledge of Content”, *Mind*, vol. 99, n° 395, 1990).

that we are not victims of the evil demon's manipulation.

Thus, according to semantic externalism, the original Cartesian evil demon hypothesis appears to be incoherent.¹¹ Yet the dispute with the sceptic hardly ends here. She can step back, give up semantic internalism, and retrench herself at a higher level. She may agree that if the content and identity of at least those beliefs which are most straightforwardly about physical objects, for instance the belief expressed by the sentence "Here is my hand", crucially depend on the believer's (or the belief's) being in proper causal connection with real physical objects, such as hands, then the Cartesian evil demon scenario cannot be coherent. But she may ask a further question about how we know that our beliefs in this, actual world have the contents we are used to ascribing to them: that they really are beliefs about physical objects such as hands, tables, trees, and so on. For the same reason we can't prove that we are not dreaming, or living in a world created by an evil demon, we are not able to show that our beliefs are caused by physical objects rather than by a spontaneous stream of dreams or by an evil demon intervening in our minds. Since we are possibly ignorant about the true causal origins of our beliefs, we can never be sure about their true contents. It turns out that semantic externalism gives ample room for the philosophical sceptic, because she may now construe Cartesian hypotheses as B-alternatives, describing us as being radically wrong about the true causal history and, consequently, the true contents of our beliefs.

Here we are faced with a second order scepticism that concerns our knowledge of the content¹² of our thoughts, particularly beliefs. But this kind of scepticism has first order scepticism as an inescapable consequence: since, as it were, we cannot exclude the possibility that the contents of our beliefs are totally different from what we suppose them to be, we cannot even meet the belief condition for knowledge of the external world.

¹¹ It is interesting to note that Descartes himself adopted an *externalist* maneuver in order to undermine the sceptical argument that appeals to the dreaming hypothesis: "Nonetheless, it must surely be admitted that the visions which come in sleep are like paintings, which must have been fashioned in the likeness of things that are real ..." (*Med.*, p. 13: 19)

¹² See Brueckner, "Semantic Answers". One additional worry for the anti-sceptical strategy grounded on semantic externalism is that it works only against some, but not all, of the ways the evil demon hypothesis can be construed: with a special scenario in which all of us have always been manipulated by an evil demon. Smart sceptic will construe the scenario in order to make it *less* vulnerable to this anti-sceptical strategy. Suppose that we have only recently become victims of an evil demon, after many years of normal embodiment and proper causal contact with real hands. In this scenario, even externalists should admit that, even now, while manipulated by the evil demon and think "Here is a hand", I do mean *here* by "here", *hand* by "hand", and so on. Thus, I can believe "Here is my hand" and would be thinking something false when I believe "I am not deceived by the evil demon".

So the sceptic wins again. Sticking to the basic sceptical strategy, she only needs one version of the Cartesian hypothesis, which will be effective enough to establish her sceptical conclusion that we don't know some things, for example that we have hands.

4. Cartesian hypotheses as J-alternatives

At this point, in discussing scepticism concerning our ordinary knowledge claims of the form "We know p ", we have dealt with the dreaming and the evil demon hypotheses as alternatives that are inconsistent with p ; the only difference was, depending on whether we choose semantic internalism or semantic externalism, whether they are construed as T- or B-alternatives. Yet they can both be slightly adjusted to be consistent with p and still remain effective. Thus, we may suppose that our dreams by some coincidence turn out to be true, or that, instead of a malicious demon, there is a mighty benevolent being (like Descartes' God) that deliberately supplies our minds with veridical perceptual experiences and true beliefs. It turns out that these sceptical alternatives, despite being construed in a way that ensures the truth of p , are still incompatible with our supposed knowledge that p : even if my belief that I am sitting here is true, I do not know whether this belief, together with all background evidence on which I base it, is merely a content of my dream or a result of someone's intervention in my mind. So specified, both alternatives obviously threaten my knowledge claim by attacking the justification condition, because the truth of my belief that I am sitting here appears to be just a lucky guess.

As far as the dreaming hypothesis is concerned, it is easy to show that it does not undermine our knowledge by precluding the truth of the propositions we claim to know, since, for instance, the proposition that I am sitting in my chair does not entail that I am not dreaming. For some reason, it might have happened that while I am sitting in my chair and working on this paper I fell asleep and started to dream that I am sitting in my chair, and thus believe it is the case; in these circumstances, no one (including myself) would be ready to credit me with the knowledge that I am sitting in my chair. More to the point, it seems that despite the fact that I meet both the truth and the belief condition, I do not know that I am sitting in my chair because I lack the appropriate justification for the belief that I am sitting in my chair. As a content of my dreams, the supposed perceptual evidence has emerged quite differently to how it is usually expected to be: in order for it to provide adequate justification for beliefs concerning the external world, it would have to have its

causal origin in the external things themselves, while as a content of a dream, it completely misses this causal connection. In spite of the fact that we might dream about some state of affairs that happens to be simultaneously taking place, the dream content can't be a source of reliable evidence since it would be the same even if the corresponding state of affairs were not to obtain. Generally speaking, whoever dreams, and in her dreams believes that p is *de facto* true, cannot be credited with knowledge that p because she simply fails to meet the justification condition for the knowledge that p .

The situation is somewhat different with the hypothesis that there is a mighty benevolent demon that controls the contents of our bodiless minds and carefully selects only those beliefs that are actually true. In these circumstances there is still some sort of causal connection between our beliefs and external things, but since it is, unbeknownst to us, mediated by the demon's decision to supply us with true beliefs only, it is entirely different from what it is supposed to be in normal circumstances. The reason why we are again reluctant to ascribe knowledge to ourselves in this context is that whether our beliefs are true or not is entirely at the disposal of the demon, and from our point of view this truth appears to be truth by coincidence.

As far as the dreaming hypothesis is concerned, it is easy to show that it does not undermine our knowledge by precluding the truth of the propositions we claim to know, since, for instance, the proposition that I am sitting in my chair does not entail that I am not dreaming. For some reason, it might have happened that while I am sitting in my chair and working on this paper I fell asleep and started to dream that I am sitting in my chair, and thus believe it is the case; in these circumstances, no one (including myself) would be ready to credit me with the knowledge that I am sitting in my chair. More to the point, it seems that despite the fact that I meet both the truth and the belief condition, I do not know that I am sitting in my chair because I lack the appropriate justification for the belief that I am sitting in my chair. As a content of my dreams, the supposed perceptual evidence has emerged quite differently to how it is usually expected to be: in order for it to provide adequate justification for beliefs concerning the external world, it would have to have its causal origin in the external things themselves, while as a content of a dream, it completely misses this causal connection. In spite of the fact that we might dream about some state of affairs that happens to be simultaneously taking place, the dream content can't be a source of reliable evidence since it would be the same even if the corresponding state of affairs were not to obtain. Generally speaking, whoever dreams, and in her dreams believes that p is *de facto* true, cannot be credited with knowledge that p because she

simply fails to meet the justification condition for the knowledge that *p*.

The situation is somewhat different with the hypothesis that there is a mighty benevolent demon that controls the contents of our bodiless minds and carefully selects only those beliefs that are actually true. In these circumstances there is still some sort of causal connection between our beliefs and external things, but since it is, unbeknownst to us, mediated by the demon's decision to supply us with true beliefs only, it is entirely different from what it is supposed to be in normal circumstances. The reason why we are again reluctant to ascribe knowledge to ourselves in this context is that whether our beliefs are true or not is entirely at the disposal of the demon, and from our point of view this truth appears to be truth by coincidence.

References

- Bourget, D. and David J. Chalmers (2013) "What Do Philosophers Believe?" (April 29) <http://philpapers.org/archive/BOUWDP.4.pdf> .
- Brueckner, A. (1990) "Scepticism about Knowledge of Content", *Mind*, vol. 99, n° 395.
- Brueckner, A. (1999) "Semantic Answers to Skepticism", in: K. DeRose and T. A. Warfield, eds. *Skepticism. A Contemporary Reader*, Oxford – New York, Oxford University Press.
- Cohen, S. (1991) "Skepticism, Relevance, and Relativity", in: Brian McLaughlin ed., *Dretske and His Critics*, Cambridge Mass., Basic Blackwell.
- Curley, E.M. (1978) *Descartes Against the Skeptics*, Oxford, Blackwell.
- Dancy, J. (1985) *An Introduction to Contemporary Epistemology*, Oxford, Blackwell.
- DeRose, K. (2000) "How Can We Know that We're Not Brains in Vats?", *The Southern Journal of Philosophy*, vol. 38, n° 51.
- DeRose, K. and Warfield, T. A. eds. (1999) *Skepticism, A Contemporary Reader*, New York – Oxford, Oxford University Press.

- Descartes, René (1985) *The Discourse on the Method*. In: *The Philosophical Writings of Descartes*, vol. I, translated by J. Cottingham, R. Stoothoff and D. Murdoch, Cambridge: CUP.
- Descartes, René (1996) *Meditations on First Philosophy, With Selections from the Objections and Replies*, translated and edited by John Cottingham, Cambridge, Cambridge University Press (f. p. 1986).
- Dretske, F. (1969) *Seeing and Knowing*, Chicago, University of Chicago Press.
- Dretske, F. (1970) “Epistemic Operators”, *Journal of Philosophy*, 67.
- Falvey, K., and Owens, J. (1994) “Externalism, Self-Knowledge, and Skepticism”, *The Philosophical Review*, vol. 103, n° 1.
- Forbes, G. (1999) “Realism and Skepticism: Brains in a Vat Reconsidered”, in: K. DeRose and T. A. Warfield, eds.
- Howthorne, J. (2004), *Knowledge and Lotteries*, Oxford, Clarendon Press.
- Lewis, D. (1996) “Elusive Knowledge”, *Australasian Journal of Philosophy*, 74.
- Putnam, H. (1981) *Reason, Truth and History*, Cambridge, Cambridge University Press.
- Unger, P. (1975) *Ignorance: A Case for Scepticism*, Oxford, Oxford University Press.
- Warfield, T. A. (1999) “A Priori Knowledge of the World: Knowing the World by Knowing Our Minds”, in: K. DeRose and T. A. Warfield, eds.
- Williamson, T. (2000) *Knowledge and its Limits*, Oxford, Oxford University Press.

**Les tendances de l'épistémologie française au tournant des
XIX^e et XX^e siècles : pragmatisme, holisme, intuitionnisme**

Léon Dumont and William James on consciousness and truth*

“The sounds of a violin improve by use in the hands of an able artist, because the fibres of the wood at last contract habits of vibration conformed to harmonic relations. This is what gives such inestimable value to instruments that have belonged to great masters.”

L. Dumont, “De l’habitude”, 1876¹

1. The importance of habit

William James’s place in the history of psychology and the history of philosophy is secured. His 1890 book *The Principles of Psychology* is a classic work in psychology and his collaboration with Charles S. Peirce from the early 1870s onwards defines him as a cofounder of pragmatism. In psychology the most significant of William James’s ideas is his view of consciousness as “one continuous stream, with no discrete ‘sensations’” in which “*qualities and things* appear to the mind in a *fringe of relations*”.² To pragmatism James contributed his well-known and controversial theory of truth. “A new opinion counts as ‘true’,” says he “just in proportion as it gratifies the individual’s desire to assimilate the novel in his experience to his beliefs in stock. It must both lean on old truth

* This paper had been presented at the conference “French Epistemology” at the University of Belgrade, Faculty of Philosophy 2013 under the title “Dumont on Habit and Will”; in the meantime I developed the points I made then which demanded the change of the title. I wish to thank the audience for the comments and questions, I am particularly grateful to the organizer Ivan Vuković for inviting me to present the paper. Also, I wish to thank Dragan Trajkovski and Timothy Williamson for useful comments during the preparation of this paper for publication. This paper is supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia through the project “Logical and epistemological foundations of science and metaphysics”, under the number 179067.

¹ A passage quoted and translated by William James in *The Principles of Psychology*, 1890. In 1952 edition, p. 69.

² William James, *The Principles of Psychology*, London, Encyclopedia Britannica, INC, 1952, p. 167, n. 2.

and grasp new fact.”³ An idea is true, claims James, if it is true *instrumentally*, that is, in so far it “works”.⁴

These two theories had diverse destinies. While James’s concept of fringes was met with appreciation by Edmund Husserl, his doctrine of truth was depreciated by Bertrand Russell. Husserl gives credit to James for the notion of horizon, which is the core notion of Husserl’s phenomenology; he declares how James “was alone in becoming aware of the phenomena of horizon under the title of ‘fringes’”.⁵ On the other hand Russell admits that he has not refuted James’s pragmatism as a whole; he says that he accepts one half of it, its method, but that he rejects the other half, its conception of truth.⁶ Russell protests that “the pragmatist theory of truth is to be condemned on the ground that it does not ‘work’.”⁷ But the half which Russell accepts – James’s pragmatic method – amounts to radical empiricism, which does not “admit into its constructions any element that is not directly experienced, nor exclude from them any element that is directly experienced.”⁸ Since we directly experience a thought with its fringes, according to radical empiricism fringes cannot be excluded: James’s pragmatic method rests on his theory of consciousness. Therefore the two halves of James’s pragmatism are his theory of truth and his theory of consciousness. And, as will be argued here, these are not two independent theses; they both stem from the same source: the theory of habit.

In developing both his theory of consciousness and his theory of truth James was profoundly influenced by the 1876 essay “De l’habitude” of a French philosopher, Léon Dumont, of which James said it was “perhaps the most philosophical account of habit yet published.”⁹ James’s contribution to the psychology of habit is well acknowledged in the sense that every paper on this topic still mentions James (though in a throat clearing way, in introductory passing remarks), but Dumont is almost never mentioned. Richard Thompson, for example, in his 2009 paper on the history of habituation notes how this term was in widespread use early in the twentieth century after the study of the

³ William James, “What Pragmatism Means” in *Pragmatism*, Cleveland, Meridian Book, The World Publishing Company, 1967, p. 52.

⁴ Cf. *Ibid.*, 49.

⁵ Edmund Husserl, *The Crisis of European Sciences and Transcendental Phenomenology*, trans. David Carr, Evanston, Northwestern University Press, 1970, p. 264.

⁶ Cf. Bertrand Russell, “James’s Conception of Truth” in *Philosophical Essays*, New York, Longmans, Green, and Co, 1910, p. 129.

⁷ *Ibid.*, 149.

⁸ William James, “A World of Pure Experience”, in: Ralph Barton Perry ed., *Essays In Radical Empiricism*, New York, Longmans, Green, and Co, 1912, p. 42.

⁹ James, *The Principles*, 69.

phenomena of habituation spread rapidly at the end of the nineteenth century. But he adds that he was unable to determine who first used the term ‘habituation’ in this context.¹⁰ So there is room for the suggestion that it might have been introduced in James’s translation of a passage from Dumont. Namely, in chapter IV of *The Principles of Psychology*, entitled “Habit”, James quotes from Dumont’s paper and “un phénomène d’habitude”, he translates as “a phenomenon of habituation.”¹¹ The reason for this neglect of Dumont’s contribution to psychology and philosophy is probably a failure to appreciate the extent of the influence he had on James, which, in turn, might have been obscured by the fact that James mentioned Dumont only once when he quoted him in the context of illustrating how philosophy of habit is “a chapter in physics rather than in physiology or psychology.”¹² Namely, Dumont claims that habit is a phenomenon of both living and dead matter, i.e. of the organic and inorganic worlds, and his view on habit had been particularly important for James because of its broadness.

In what follows I am going to provide evidence of Dumont’s influence on James. By representing James’s views on consciousness and truth as extensions of Dumont’s views on the habitual nature of will and the habitual nature of belief, I wish to show in which directions Dumont’s positions might have developed since, as was pointed out by his contemporary Joseph Delboeuf, “his thought didn’t and couldn’t reach the full maturity.”¹³ Namely, Dumont who was born on February 5th 1837 died on January 17th 1877 before turning 40, “still searching for his fixed point, which he probably would have found eventually.”¹⁴ This “independent, serious and original spirit,” continues Delboeuf, didn’t belong to any school, he was eclectic, honest and conscientious and his criticism was always penetrating.¹⁵ Although eclectic in his interests, Dumont didn’t belong to the school of eclecticism of Victor Cousin which was dominant in his time. Rather, Dumont’s eclecticism was similar to James’s, based on his vast capacities and desire to learn and travel, his wide interests and good knowledge of languages. Dumont was under the influence of Scottish philosophy due to the philosophy professor Mr. Mathoré in his home town of Valenciennes where he had studied. Delboeuf described Mathoré as a modest and

¹⁰ Cf. Richard Frederick Thompson, “Habituation: A History”, *Neurobiology of Learning and Memory*, vol. 92, 2009, p. 127.

¹¹ James, *The Principles*, p. 69.

¹² *Ibid.*, pp. 68-69.

¹³ Joseph Delboeuf, “Léon Dumont et son oeuvre philosophique”, *Revue Philosophique de la France et de l’étranger*, vol. 6, 1877, p. 622.

¹⁴ *Ibid.*, p. 623.

¹⁵ *Ibid.*

educated man, who, particularly appreciated Thomas Reid and Dugald Stewart.¹⁶ This explains the spirit close to pragmatism of Dumont's most mature paper "De l'habitude" that was recognized by William James who, like Charles Peirce, also owed much of his philosophical standpoint to the Scottish school of philosophy.

2. The broadness of habit

At the beginning of the nineteenth century it was the French physiologist Xavier Bichat who introduced habit into the focus of philosophical research. Unlike Dumont and James, Bichat saw habit as a very exclusive phenomenon. He excluded habit not only from the realms of dead matter and plants, but even from the vegetative part of animal life. This part of life – which is common to plants and animals – Bichat calls organic life. It comprises functions like circulation, respiration, exhalation, absorption, nutrition, and secretion. None of these, Bichat claims, are regulated by habit.¹⁷ On the other hand, functions such as perception of the surroundings, reflection on sensations, and voluntary movements Bichat relates only to the animal part of life, claiming that everything in animal life is regulated by habit.¹⁸ So, habit was considered by him one of the demarcating characteristics between the organic and animal spheres of life.

Questioning this, Dumont gives the example of *Mimosa pudica*, a plant that closes its leaves every evening. If we perform an experiment and put a *Mimosa pudica* in some dark place and illuminate it during the nights, the *Mimosa* will, for some period of time, continue to open its leaves in the daytime and close them in the evenings. But in the long run it will change its habits and start to sleep in the daytime and open its leaves at night.¹⁹ This plant also closes its leaves when it is touched, so the physiologist Jules-Bernard Luys compares the sensitivity of *Mimosa pudica* with that of an animal: it is "like an animal," he says "it feels and reacts on the contact of the lightest touch."²⁰ Departing from Bichat, Dumont and Luys maintain that habit connects the organic and animal parts of life. If habit is not a demarcating but a uniting line between the two spheres of life, Dumont and

¹⁶ *Ibid.*, p. 612.

¹⁷ Marie-François-Xavier Bichat, *Physiological Researches upon Life and Death*, trans. Tobias Watkins, Philadelphia: Smith & Maxwell, 1809, p. 40.

¹⁸ *Ibid.*, 34.

¹⁹ Cf. Léon Dumont, "De l'habitude," *Revue philosophique de la France et de l'étranger*, vol. 1, 1876, p. 330.

²⁰ Jules Bernard Luys, *The Brain and Its Functions*, New York, D. Appleton and Company, 1882, p. 86.

James are justified in seeing living beings as systems of habits. Living individuals, claims Dumont, are “systems of habits interacting with each other in a constantly variable equilibrium.”²¹ James, in a similar vein, describes human creatures as walking bundles of habits. He says: “when we look at living creatures from an outward point of view, one of the first things that strike us is that they are bundles of habits.”²²

Some properties are generally understood as defining attributes of habits. These features of habituation are, among others, overcoming resistance, becoming easier on repetition, saving trouble, and improving functioning. “A lock works better,” says Dumont, “after being used some time; at the outset more force was required to overcome certain roughnesses in the mechanism. The overcoming of their resistance is a phenomenon of habituation. It costs less trouble to fold a piece of paper when it has been folded already. This saving of trouble is due to the essential nature of habit, which brings it about that, to reproduce the effect, a lesser amount of the outward cause is required.”²³ Does the fact that these features are acquired by a lock or a piece of paper entail that inorganic things can also acquire habits? Disputing the assumption that inorganic things can acquire habits, Dumont’s contemporary Joseph John Murphy says: “The definition of habit, and its primary law, is that all vital actions tend to repeat themselves; or, if they are not such as can repeat themselves, they tend to become easier on repetition. It may appear that this law is in no way peculiar to the actions of living beings: for there are many inorganic actions that tend to repeat themselves, and to become easier on repetition. For instance: flowing water generally makes a channel for itself, and tends to flow afterwards in the same channel; and if a piece of paper has been once folded, it is easier to fold it again in the same folds than in new ones. But there is a fundamental difference,” thinks Murphy, since in the cases described “the direction of action is determined by *mere change of form*: the water tends to flow in the channels, because their form is suitable; the paper tends to lie in particular folds, because it has acquired their form.”²⁴ Dumont agrees that habit in inorganic matter is not a change in a substance, but he disagrees that it is a mere change of form; the changes in inorganic matter, he asserts, are new *habits of cohesion*. He says: “Everyone knows how a garment, after having been worn a certain time, clings to the shape of the body better than when it was new; there has been a change

²¹ Dumont, *op. cit.*, p. 321, when not specified otherwise, paraphrases and translations from Dumont’s essay are mine.

²² James, *The Principles*, p. 68.

²³ Dumont, *op. cit.*, 323-324, quoted and translated by James in *The Principles*, p. 69.

²⁴ Joseph John Murphy, *Habit and Intelligence* (London: Macmillan and CO., 1869), vol. 1, p. 169.

in the tissue, and this change is a *new habit of cohesion*.²⁵ Dumont's argument that cohesion is of the nature of habit is based on the distinction between adaptations and habits. Habits are the products of the adjusting reactions of the structure which has been exposed to some influence. That is, habits are *states* that are the results of the *events* that are processes of adaptation.²⁶ Since cohesion is defined as "a state in which all parts of something fit together well and form a united whole,"²⁷ it can be said that cohesion is a state that is reached through the processes of adaptation of the molecules of a thing. Namely, inside one and the same inorganic thing there is an attractive force between its molecules which interact with each other in, it can be said, 'constantly variable equilibrium'. This equilibrium is cohesion. Or, as Dumont defines it, cohesion is the balance achieved "from myriads of elementary movements limiting each other."²⁸ As a state of equilibrium, cohesion is a habit of a kind. Therefore, the conclusion that inorganic things acquire habits follows from the premise that cohesions in inorganic things are habits. Dumont finds crystals to be good examples of this. For instance, common salt crystallizes in pure water in cubical form, but if the water contains certain acids, the angles make truncated cubes.²⁹ It is true that crystals change their form in different solutions. But Dumont's point is that these changes are not mere changes in form since they consist in altering the mode of the crystal's growth. Namely, the various solutions are not just different conditions, but are different specific nutritive conditions under which the chemical composition – the substance – does not change, but the structure of binding – the phenomenon – changes. But it is striking that while we do not see problems in talk about the habits of cohesion of crystals, we do hesitate, for example, to say that a piece of paper or a door is habitually such and such. We are not willing to say of a piece of clay that in the hands of an able sculptor it acquires the habit of being *The Thinker* (*Le Penseur*). It does not sound equally weird to talk of habits of cohesion of clay. But the fact that some implications do not sound familiar is not a proof that a thesis that entails them is not valid. That a piece of clay acquires a habit follows from the plasticity of this material. Namely, James claims that habits in general are due to "plasticity of materials to outward agents".³⁰ Plasticity, says James, "means the possession of a structure weak

²⁵ Dumont, *op. cit.*, 323, quoted and translated by James in *The Principles*, p. 69.

²⁶ *Cf. Ibid.*, 336-337.

²⁷ *Essential English Dictionary*, London, Collins, 1989.

²⁸ Dumont, *op. cit.*, p.325.

²⁹ *Ibid.*, p. 324.

³⁰ James, *The Principles*, p. 70.

enough to yield to an influence, but strong enough not to yield all at once. Each relatively stable phase of equilibrium in such a structure is marked by what we may call a new set of habits.”³¹ And, if without hesitation we say that a lump of clay due to its plasticity turned into *Le Penseur*, why couldn’t we then say that this equilibrium is a new set of the lump’s habits? If it does not seem problematic to say that plasticity is a feature of inorganic matter and that plasticity is the root of habits, why should it be problematic to say that inorganic matter acquires habits?

Giving a final reason why inorganic things do not acquire habits, Murphy says that the tendencies acquired by a piece of paper or a stream of water will be utterly lost if the folds are taken out of the paper by hot pressure or the channels are filled up.³² But this exactly proves Dumont’s view that the greatest enemy of a habit is another habit.³³ That is, in order to lose a habit one should acquire an opposite habit. Dumont notices how the best cure for morbid repetitive habits is a complete change of life style.³⁴ Or, as James advises: “to conquer undesirable tendencies in ourselves, we must ... go through the outward motions of those contrary dispositions we prefer to cultivate.”³⁵ So, the fact that a piece of paper lost the traces of previous folds does not prove that what it had was not a habit, only that it was lost by acquiring a new one.

Finally, since living beings, plants and inorganic things all acquire habits, Dumont concludes that “habit is a universal fact,”³⁶ and that the whole of “nature is a result of habits.”³⁷ James, along similar lines, says: “The laws of nature are nothing but the immutable habits which the different elementary sorts of matter follow in their actions and reactions upon each other.”³⁸

3. Habit and consciousness

An act done out of habit is usually related to an unconscious act. We say: “I wasn’t aware that I turned left; I did it out of habit,” as if a habitual act excludes awareness. There

³¹ *Ibid.*, p. 68.

³² Murphy, *Habit and Intelligence*, p. 169.

³³ *Cf.* Dumont, *op. cit.*, 363.

³⁴ *Ibid.*, p. 352.

³⁵ William James, “What is an emotion?” in *The Emotions*, ed. K. Dunlop, Baltimore, Williams & Wilkins Company, 1922, p. 22.

³⁶ Dumont, *op. cit.*, p. 322.

³⁷ *Ibid.*, p. 321.

³⁸ James, *The Principles*, p. 68.

are two main reasons for this. First, habit, in general, makes it possible to do things with less conscious effort. And second, as Dumont notices: “very often habit makes possible the elimination of certain phenomena which originally served as intermediaries between two acts. Where there was an absence of these phenomena, it was believed that there it was merely unconscious.”³⁹ That no unconscious ideational intermediary is needed follows from the fact that inorganic matter acquires habits as well. Dumont argues from the following analogy. The hand of a worker might be needed to adjust two parts of a mechanism, let’s say two fly-wheels. But when these parts of the mechanism adapt to each other, the finger used to adjust them is not needed anymore. Analogously, if an adaptation of a reflex character is established between a perception and a gesture, then there is no need for intelligence to intervene.⁴⁰ For instance, a musician learns to play the piano first with the help of his intellect and his will which establishes a constant adaptation between the muscle movement of a finger and perception of the notes. But once the adaptation is produced (when he acquired a habit of performing well) the phenomenon of sight is sufficient to cause a movement of the finger without the need for an intermediate intervention of intelligence and control.⁴¹ Support for this claim of Dumont’s can be found in *The Principles*, where James quotes Georg Schneider and William Carpenter. Schneider describes a beginner at the piano who “not only moves his finger up and down in order to depress the key, he moves the whole hand, the forearm and even the entire body, especially moving its least rigid part, the head, as if he would press down the key with that organ too.”⁴² Carpenter, on his side, describes an experienced musical performer to whom the accustomed sequence of movements is “directly prompted by the sight of the notes.”⁴³ This difference between a beginner and an expert shows how something that was needed and present during the adaptations that led to a habit’s acquisition need not remain present (either explicitly or implicitly) in the manifestation of the habit – just as the finger used for adapting fly-wheels to each other, does not remain latently present after the adjustment is established. But in the case of the skills described, although intermediary ideational elements are not explicitly present, we can ask whether they remain present implicitly. Dumont argues against philosophers who

³⁹ Dumont, *op. cit.*, p. 328.

⁴⁰ *Ibid.*, p. 328-329.

⁴¹ *Ibid.*, p. 360.

⁴² From G. H. Schneider, *Der menschliche Wille vom Standpunkte der neueren Entwicklungstheorien*, 1882, free translation by James, *The Principles*, p. 73.

⁴³ From W. B. Carpenter, *Principles of Mental Physiology*, 1874, quoted in James, *The Principles*, p. 76.

believed that if some idea had a role in establishing a connection between two other ideas, it still exists in the mind when not necessary anymore, that it remains unconsciously present. For example, if an idea A (say, of watching a movie) suggests an idea B (of going to a cinema), and B suggests C (of reading a cinema guide), due to repetitions there occurs coexistence between A and C, which in the long run results in A alone suggesting C. From the original chain of ideas A-B-C, through habituation the intermediary act B, Dumont claims, when it becomes unnecessary can be omitted. It becomes completely absent; it is not present in any latent, implicit, unconscious way.

This phenomenon is governed by two laws. The first one is the principle of habit called the “diminution of causality” (*le principe de la diminution de causalité*)⁴⁴ which George Stout attributed to Dumont. A habitual process is “capable of being set in action by a slighter cue. If on the first occasion it has occurred by way of reaction to a complex stimulus *a b c d*, it may, after frequent repetition, find sufficient prompting in the presence of *a b c* or *a b*, or even of *a* alone.”⁴⁵ This principle explains why in a causal chain of events A-B-C, the link B becomes superfluous. The second law involved is stated by James as a general law of consciousness; it is *the principle of parsimony* in consciousness, according to which “consciousness seems to desert all processes where it can no longer be of any use.”⁴⁶ This principle is an instance of the first one for it assumes that the nervous system functions as a system of habits and applies the principle of diminution of causality to the chains of thoughts. It should be noted that James prefers to talk about thoughts instead of consciousness. He says he uses the word ‘thinking’ “for every form of consciousness indiscriminately”.⁴⁷ This choice is due to James’s general position regarding consciousness, which is that consciousness does not exist as an entity, but as a function.⁴⁸ Also, strictly speaking, James wouldn’t speak of chains of thoughts either. Namely, dealing with the question how consciousness appears to itself, James notes how it does not resemble a chain or a train for it has no joints, it simply flows.⁴⁹ It resembles water in a way vividly depicted by Dumont. James quotes Dumont: “water, in flowing, hollows out for itself a channel, which grows broader and deeper; and, after having ceased to flow, it resumes, when it flows again, the path traced by it before. Just so, the

⁴⁴ Dumont, *op. cit.*, p. 354.

⁴⁵ George Frederick Stout, *Analytic Psychology*, London: George Allen & Unwin LTD, 1918, p. 263.

⁴⁶ William James, *The Feeling of Effort*, Boston: Boston Society of Natural History, 1880, p. 5.

⁴⁷ James, *The Principles*, p. 146.

⁴⁸ William James, “Does ‘Consciousness’ Exist?” in *Essays In Radical Empiricism*, ed. Ralph Barton Perry (New York: Longmans, Green, and Co, 1912), p. 3.

⁴⁹ James, *The Principles*, p. 155.

impressions of outer objects fashion for themselves in the nervous system more and more appropriate paths and these vital phenomena recur under similar excitements from without, when they have been interrupted a certain time.”⁵⁰ It is possible that Dumont’s description inspired James to coin his famous metaphor of the stream of thought.

In favor of the claim that unconsciousness is not a necessary trait of a habitual act is also the fact that such an act can be exercised in the presence of consciousness. Moreover, some important faculties that we exercise with awareness are of the nature of habit. For example, both Dumont and James think that memory is based on habit. James says that “the cause both of retention and of recollection is the law of habit in the nervous system.”⁵¹ Dumont quotes Gassendi who, opposing the view that memory is like a storehouse, famously compared memory and habit with a piece of paper which can be more easily folded if it had been folded before. Memory’s functioning is described as based on hitting the right fold.⁵² This comparison, according to Dumont, gives a good idea of acquired habits: their exercise comes back easily, without a period of adjusting and remembering.⁵³

Although the claim that dead matter acquires habits seems to be in accord with the thesis that habits are unconscious, involuntary phenomena, the crucial proof for the claim that unconsciousness is not a feature of habit comes from the fact that habit is a phenomenon in the inorganic world. If, of a door which is always closed, we can say that it is habitually closed, we will not claim, says Dumont, that it abides latently closed when it is open. Or, of a piece of paper that is habitually folded in one way, when it is folded differently it makes no sense to say that it remains latently folded in another way.⁵⁴ Analogously, an intermediary idea that is not consciously present in a chain of ideas is not present unconsciously. Dumont concludes: “When we no longer think of the idea, it no longer exists even in a latent state.”⁵⁵ Finally, if there is no sense to saying that something is conscious, then there is no sense to saying that it is unconscious. Dead matter is out of the realm of consciousness so, in general, the features of dead matter cannot be described as being unconscious.

⁵⁰ Dumont, *op. cit.*, pp. 323-324 (quoted and translated by James, *The Principles*, p. 69).

⁵¹ James, *The Principles*, p. 427.

⁵² Cf. George Sidney Brett, *Philosophy of Gassendi*, London: Macmillan and Co. Limited, 1908, pp. 133-134.

⁵³ Cf. Dumont, *op. cit.*, p. 348.

⁵⁴ *Ibid.*

⁵⁵ *Ibid.*, p. 351.

4. The habitual nature of the will and the fringed structure of consciousness

We usually relate the habitual nature of an action with its being involuntary. We say: “I didn’t want to turn left, I did it out of habit”, as if habit excluded voluntariness. Dumont claims that it is not a necessary trait of habit to be involuntary. Plants exhibit some degree of sensitivity as they react to touch like *Mimosa pudica*, but they do not perform any voluntary movements whatsoever. Since motor centres are responsible for voluntary acts, it follows that motor centres are not an essential factor for exhibiting habits. If there is no sense to speak of voluntary acts of plants, it makes no sense to speak of involuntary ones either. Therefore, habitual acts are not necessarily involuntary. In fact, some habitual acts appear automatic but are at the same time wanted. Dumont gives the example of an employee who rises in the morning, gets dressed and leaves the house without the idea of where he needs to go, but when a friend meets him and asks why he got up so early, he replies that he wanted to go to his office.⁵⁶ So the fact that a worker goes to work out of habit does not exclude him wanting so. Moreover, for Dumont “the will is nothing but a manifestation of a habit that had been previously acquired.”⁵⁷ This follows from the two elements of the will Dumont distinguished.

The first element of the will is that the idea of a goal may suggest the idea of an act. This is possible due to habit because it allows intermediaries between the idea of an act and the idea of its goal to disappear from the original chain of ideas that relates a particular action to its end. For example, as Dumont explains, we know from experience that an act A causes an event B which causes a result C. Then, under the influence of habit, the act A and the goal C become adapted to each other so that B is not needed anymore in the chain A-B-C. The coexistence between A and C becomes constant, which enables us, according to Dumont, to reverse the direction of the suggestion between A and C, so that the power of suggestion goes from the idea of the goal C to the idea of the act A.

The second element of the will, according to Dumont, is that the idea of an act can suggest that act. This is also possible due to habit. Namely, we are in the habit of noticing what we are doing. Our acts are usually accompanied by our representations of these acts.

⁵⁶ Cf., Dumont, *op. cit.*, p. 333.

⁵⁷ *Ibid.*, p. 331.

For example an act A is followed by its representation Ra, and through habit these two become adapted to each other, so that the direction of suggestion is reversed: the representation Ra suggests the act A. If this idea is strong enough, it can even cause the act A.⁵⁸ For example, the idea of walking simply makes us walk, claims Dumont, and he defines the will as the causation of an act with the help of an idea of its consequences.⁵⁹ Similarly James says: “every representation of a motion awakens the actual motion which is its object,” but he adds “unless inhibited by some antagonistic representation simultaneously present to the mind.”⁶⁰ If it happens that two ideas of movement conflict, the one that prevails triggers an action. The proof of this thesis is based on a thesis, argued for by James in *The Feeling of Effort* (1880), that we can feel only afferent impulses coming from receptors and therefore our will cannot consist in our feeling of the force exerted but we can only compare our ideas of certain actions, and these ideas consist in the sets of afferent impulses that determine a certain act. The feeling of muscular energy is “a complex afferent sensation.”⁶¹ Since we cannot feel efferent impulses, we are comparing not the ideas of performing actions but the ideas of the sensitive results of the actions. Or as Dumont says: “Will is only a consequence of anticipating the effects of some action we had performed under any excitation.”⁶²

How then can willed movements actually be accomplished? At first by chance. For instance, a child with a goal in view makes random movements until she accidentally finds the right one. “When a particular movement, having once occurred in a random, reflex, or involuntary way,” explains James, “then the movement can be desired again, proposed as an end, and deliberately willed. But it is impossible to see how it could be willed before.”⁶³ “Instead of doing what we want,” concludes Dumont “we want what we do.”⁶⁴ Even when we want to do something we have never done, this act is a new combination of basic acts each of which depends on habit. So every voluntary act involves habit; every act we want demands an act we are used to.⁶⁵

For both Dumont and James, the mind is in constant change. “Every thought we have of a given fact is, strictly speaking unique,” James says, it “only bears a resemblance

⁵⁸ *Ibid.*, pp. 330-334.

⁵⁹ *Ibid.*, p. 330.

⁶⁰ James, *The Feeling of Effort*, p. 17.

⁶¹ *Ibid.*, p. 4.

⁶² Dumont, *op. cit.*, p. 333.

⁶³ James, *The Principles*, p. 768.

⁶⁴ Dumont, *op. cit.*, p. 333.

⁶⁵ *Cf. Ibid.*, p. 334.

of kind with our other thoughts of the same fact.”⁶⁶ We have the illusion of having the same thoughts and feelings because the same object can appear twice, but “an unmodified [the same] feeling is an impossibility” says James because every feeling produces a brain modification and to “every brain-modification, however small, must correspond a change of equal amount in the feeling which the brain subserves.”⁶⁷ In support of the same claim Dumont quotes Étienne de Condillac with whom he agrees that all ideas are like sensations “the modifications of the soul. They exist as they alter it and they no longer exist as they stop changing it.”⁶⁸ The assimilation of ideas and sensations is accepted by James as well. As a general term comprising both ideas and sensations James uses ‘thought’, and no thought or sensation can appear twice since thought is in constant change.⁶⁹

The principle of change is habit. Habit is a change that prevents destruction since it is a balance established in response to some external influence. The constantly changing sensations appear in the fringes of relations due, as I am going to argue, to the habitual nature of the will. I particularly wish to point out how James’s understanding of consciousness as having a fringed structure follows from two elements of the will as they are analyzed by Dumont together with James’s thesis that we can feel only afferent impulses. The fringe structure means that no sensation can appear to consciousness isolated, but is always related to something else in our experience, and it always comes with a fringe. When there is a chain of actions from A to the goal Z consisting of acts A, B, C,...Y, Z then through habit the original chain A-B-C-...-Y-Z becomes shorter and shorter, until finally an act A becomes enough to yield Z. Through repetition, the elements of the pair A-Z become coexistent which makes it possible, according to Dumont, to reverse the order of suggestion, so that an idea of a goal Z, due to the first element of the will, suggests an idea of an act A. An act A will be actually prompted due to the habit established between doing A and noticing performing A which can be reversed due to the second element of the will.

James showed how our idea of an action is determined by a set of afferent impulses related to the realization of it. This means that an idea of every action (say A) is determined by a corresponding sensation (say *a*). So a chain of the ideas of actions from

⁶⁶ James, *The Principles*, p. 151.

⁶⁷ *Ibid.*

⁶⁸ Dumont, *op. cit.*, p. 350.

⁶⁹ Cf. James, *The Principles*, p. 148-150.

A, B, C, ... to Z corresponds to the chain consisting of *a, b, c, ..., z*. “In an habitual action,” claims James “mere sensation is a sufficient guide.”⁷⁰ James gives an example in which A, B, C, D, E, F, G represent a habitual chain of muscular contractions and *a, b, c, d, e, f, g* stand for the respective sensations which these contractions excite in us when they are successively performed.⁷¹ “If an act require for its execution a chain, A, B, C, D, E, F, F, G, etc., of successive nervous events,” says James, “then in the first performances of the action the conscious will must choose each of these events from a number of wrong alternatives that tend to present themselves; but habit soon brings it about that each event calls up its own appropriate successor without any alternative offering itself, and without any reference to the conscious will, until at last the whole chain, A, B, C, D, E, F, G, rattles itself off as soon as A occurs, just as if A and the rest of the chain were fused into a continuous stream.”⁷² We may conclude that in the case of a habitual action, for which mere sensation is enough, this continuous stream is a stream of sensations, in which no sensation appears in isolation but with a fringe.

5. The voluntary nature of belief and the instrumental theory of truth

While volition is the idea of a movement that prevailed, an idea, in general, that prevails is a belief. A controversial claim of James’s is that a belief that prevails is a truth. Let us see how it developed.

The phenomena of habit in organic life are possible due to the plasticity of organic matter. Due to the plasticity of their structures, living beings develop and change while maintaining their integrity. The most plastic of organic matter is the brain, and one of the stable phases of the brain’s equilibrium is belief. This justifies Dumont in his claim that beliefs are our habits of thought. Beliefs are habits of our ideas in the sense that they are the results of processes of adaptation our ideas have been going through. About these processes Dumont says: “An idea (a thought) comes to us for the first time, it is contrary to our beliefs, to our habits of thought. We declare it false, but if it returns supported by

⁷⁰ *Ibid.* p. 75.

⁷¹ *Ibid.*

⁷² *Ibid.*, p. 74-75.

new arguments, step by step the power of ideas which resisted it is destroyed; we begin to find the idea (thought) sustainable.”⁷³

After a new idea is, step by step, supported by arguments “contrary ideas gradually disappear and are replaced by others which are likely to adapt to the new arrival; therefore, the idea is *true for us*,” says Dumont, “it is rooted, we can no longer dispense with it; we are somehow accustomed to the poison and these contrary ideas become *untrue to us*.”⁷⁴ This claim of Dumont’s that an idea is true or untrue for us is kindred to James’s claim that truth depends on a subject’s desire to assimilate a new belief to her beliefs in stock.

Acquiring new beliefs Dumont compares with the action of poison. The poison “kills in large doses,” he says, but “in graded doses and often repeated, it produces habituation.” This is the case because then “organs have enough time to get used to partial destruction, to suppress resistance; and when the habit is complete, the poison is not poisonous anymore.”⁷⁵ James takes Dumont’s lead when he says: “The individual has a stock of old opinions already, but he meets a new experience that puts them to a strain. ...the result is an inward trouble to which his mind till then had been a stranger, and from which he seeks to escape by modifying his previous mass of opinions. He saves as much of it as he can, for in this matter of belief we are all extreme conservatives.”⁷⁶ People stick to their beliefs long after they have forgotten the reasons for them. Habit keeps together people who have stopped liking each other, because it is difficult to get rid of habits even when they are no longer pleasant.

Taking a belief as a habit is one thing, but it is much less plausible to claim that truth is subjective. In defending this view, James claims that the truth is also a habit. “Truth happens to an idea,”⁷⁷ says James it “does not exist, it holds or obtains.”⁷⁸ Ideas go through processes of adaptation, getting into stable phases called beliefs, but beliefs go through adaptations as well – they adapt to each other in order to reach stable states that are called true beliefs. “Truth becomes”, says James “a habit of certain of our ideas and beliefs in their intervals of rest from their verifying activities.”⁷⁹ Although here James speaks of truth as a habit of ideas and beliefs, he does not assimilate beliefs and ideas.

⁷³ Dumont, *op. cit.*, p. 356.

⁷⁴ *Ibid.*

⁷⁵ *Ibid.*

⁷⁶ James, “What Pragmatism Means”, p. 50.

⁷⁷ William James, “Pragmatism’s Conception of Truth” in *Pragmatism*, Cleveland, Meridian Book, The World Publishing Company, 1967, p. 133.

⁷⁸ *Ibid.*, p.147.

⁷⁹ *Ibid.*, p. 145.

Every belief is an idea, but not every idea is a belief. For James, a belief is an idea adopted by the mind. “The mind’s espousal of the idea is called Belief,” says James.⁸⁰ Plasticity is also a trait that enables beliefs to acquire habits as well. James’s claim precisely stated is that a belief is a stable state of ideas, therefore belief is a habit of ideas. That James’s theory of truth is a natural extension of his theory of habit is manifest in his claim that even the oldest truths are plastic.⁸¹ Namely, truth is a stable state of beliefs, it is the result of the process of the adaptation of beliefs, and hence truth is a habit of beliefs.⁸² As habits, truths are hard to change - loyalty to older truths is the first principle⁸³ – but as constantly variable equilibriums they are vulnerable and liable to a change, which is proved, observes James, by the transformation of logical and mathematical ideas.⁸⁴

Finally, how can it be reasonable to claim that a subjective feeling of some idea being in coherence with some others is a better guide to their validity than trains of reasoning a subject can offer in support of it? The answer can be found in Dumont’s paper. Dumont quotes Pierre Maine de Biran who says “When a chain of reasoning has become very familiar to us by frequent repetition, we neglect the operation which motivated it in beginning and by dint of neglecting them or of going over them quickly, we end by forgetting them, misunderstanding them, and considering them absolutely useless. This is what warrants so many ellipses in the forms of reasoning as in those of usual and familiar discourse.”⁸⁵ In other words, what a subject feels is in harmony with his other beliefs is not just his subjective feeling, but is based on a firm objective network of neural paths traced previously. The more stable this neural foundation of one’s beliefs is, the more hidden from the subject are the abstract chains of reasoning that represent the justificational grounds for them.

To conclude, the same elements that make the will possible make these ellipses and inversions in our inferences possible, even inevitable eventually. This explains why James says that will and belief are just “two names for one and the same psychological phenomenon.”⁸⁶ That is, what mind does in both cases, says he, is the same: “It takes an image, and says, ‘so far as I am concerned, let this stand; let it be real for me.’”⁸⁷

⁸⁰ James, *The Feeling of Effort*, p. 26.

⁸¹ James, “What Pragmatism Means”, p. 53.

⁸² James, “Pragmatism’s Conception of Truth”, p. 145.

⁸³ Cf. James, “What Pragmatism Means”, p. 51.

⁸⁴ *Ibid.*, p. 53.

⁸⁵ Pierre Maine de Biran, *The Influence of Habit on the Faculty of Thinking*, trans. Margaret Donaldson Boehm, Baltimore: Williams & Wilkins, 1929, p. 203. Quoted in Dumont, *op. cit.*, p. 362.

⁸⁶ James, *The Principles*, p. 661.

⁸⁷ James, *The Feeling of Effort*, p. 26.

Summarizing the results of the previous discussions, one can say that Dumont's theory of universality of habit links James's fringed theory of consciousness and instrumental theory of truth. For the universality of habit Dumont gives an ingenious argument based on the nature of cohesion. The universality of habit makes possible the exclusion of mediating steps in the habitual chains of actions, and importantly it facilitates the inversion between the final and initial members of these chains. This is the key step for the development of the will, as well as for treating beliefs as habits of our thoughts. The fringed structure of the consciousness follows from the habitual nature of the will and the fact that we can only feel afferent impulses. The instrumental theory of truth follows, according to James, from the voluntaristic nature of belief and the fact that plasticity is the root of habits.

We can conclude that James's radical empiricism and his pragmatist theory of truth are not two independent theses – that a distinctive feature of James's pragmatism is the close connection between his theory of consciousness and his theory of truth through his theory of habit.

On the Duhem-Quine thesis*

The first thing that comes to mind upon mentioning the names of Duhem (*Pierre Duhem*) and Quine (*Willard Van Orman Quine*) together is the thesis that in the philosophical literature goes under the name of *Duhem-Quine Thesis (DQT)* or the *Thesis of Underdetermination*. Although agreeing on many points, there are strong arguments which suggest that the two philosophers did not hold mutually equivalent theses, and that Duhem would be reluctant to accept what we could dub as *Quine's Thesis (QT)*: any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system.¹ On the other hand, there is a common core of the thesis they both shared. We shall refer to this common core as the *Duhem's Holistic Thesis (DT)*: *empirical statements are so intertwined that we cannot talk about refutation or confirmation of particular empirical statements alone*.

Our aim is to show that we should ascribe to Duhem only the weaker, **DT** and that the theses **DT** and **QT** are by no means identical. Hence, it will follow that these theses are essentially different from **DQT** also. As already mentioned, the thesis **DQT** also goes by the name of *Thesis of Underdetermination*. In short, this thesis asserts the possibility of incompatible, empirically equivalent theories that explain and are confirmed to the same degree by all possible empirical evidence. Therefore, empirical evidence cannot be counted as a basis for favoring one over the other. In what follows it will become clear that neither **DT** nor **QT** (as formulated in “Two Dogmas of Empiricism”) are concerned with the possibility of empirically equivalent theories. These are first and foremost holistic theses which, along with some additional assumptions, imply **DQT**.

In his most important work *The Aim and Structure of Physical Theory*, Duhem's aim was to explain the specific structure, role and aims of physical theories by critically

* This paper was supported by the project “Dynamic Systems in nature and society: philosophical and empirical aspects” (evidence #179041) financed by the Ministry of Education, science and technological development of the Republic of Serbia.

¹ W. V. Quine, “Two Dogmas of Empiricism”, in: W. V. Quine, *Quintessence*, ed. R. F. Gibson, Jr., Cambridge Mass., Belknap Press of Harvard University Press, 2004, p. 51.

examining realism that was prominent amongst philosophers and physicists of the late XIX century. This realism was founded on a belief that science is in a constant search for truth, so that, although the theory can never be confirmed, we could hope that by eliminating rival hypotheses we will isolate the one that is true in the end. The procedure which allows us to reject one hypothesis while confirming the other is called *the crucial experiment*.

By rejecting these realistic assumptions of a theory reflecting reality, Duhem is favoring the instrumentalist view on theories as economical representation of physical laws and their classification.² If the sole aim of a physical theory is to represent and classify experimental laws, then the only possible way we could measure its success is by comparing its consequences with these experimental laws. Scientific laws are neither true nor false – they merely approximate. Since they are only symbolizing nature, we cannot expect them to provide the ultimate picture of reality. Apart from demands for empirical confirmation and simplicity, Duhem also insisted that the theory be as minimalistic as possible, i.e. to shun from explanations that which concerns the unobservable aspects of natural phenomena.

Duhem modified two aspects of classical empiricism (and in this respect his position can be likened to Quine's³): while traditional empiricism was primarily concerned with empirical basis of natural science, Duhem is more focused on theoretical interpretation of that basis. In doing this, he also rejected the inductivist picture of science which was characteristic for the period he was working in. Proponents of inductivism, most notably Newton, held that science should shun from hypotheses which concern unobservable entities and processes and accept only laws arrived at by the use of induction. For Duhem, on the other hand, this presents an unacceptable description of what scientists should be doing. Namely, even though there are circumstances in which scientists should rely on inductive methods, they are not characteristic of all sciences, especially physics.

Duhem pointed out that what is crucial for science is not the exact nature of what we directly perceive, but rather the interpretation we are ascribing to it, i.e. theoretical notions as opposed to raw data which stand for that which we directly perceive. Moreover,

² Pierre Duhem, *The Aim and Structure of Physical Theory*, Princeton, Princeton University Press, 1991, p. 23.

³ Mary Hesse, "Duhem, Quine and a New Empiricism", in: Sandra Harding ed., *Can Theories be Refuted*, Dordrecht-Boston, D. Reidel Publishing Company, 1975, pp. 186-188.

the relation between these is not one-to-one, since to a single object of perception there could correspond countless more or less accurate idealizations. Therefore, if we try to, for example, fix the notion of temperature or weight, we shall recognize the need for some kind of theoretical interpretation – experience alone will not be enough. The basis of science consists in a set of theoretical notions used in interpreting experience. Since their relation to experience is not completely clear, we cannot be certain they provide us with a firm empirical basis for science.

According to Duhem, theoretical terms are not susceptible to operational or explicit definition via observational terms, rather they are in most cases introduced into science via scientific theories themselves. Therefore, theoretical terms are more often than not implicitly defined by the theory. However, since from a theory we are able to derive statements which do not contain theoretical terms and which are used for testing it, we can speak about a theory having empirical content and being susceptible to confirmation or refutation. Since theoretical statements contain theoretical terms which are only partially empirically interpreted, these cannot be tested in isolation. What Duhem concluded was that theories are measured against experience as a whole.

Inside physical theories, for instance, we find terms with unobservable referents: mass, temperature etc. These terms are operationally defined – we define something by the procedure verifying its existence, duration or quantity. For example, weight is defined in the following way: *weight* of an object is the number which appears when we put the said object on a scale. It should be clear that the meaning of these terms is dependent on the methods of measurement, but it is also dependent on the theory of the measuring instruments. These, on the other hand, are implicitly defined by a theory of measuring instruments. Every application of an instrument involves stipulated and causal elements: causal elements are reflected in the assumption that during each measuring process there is a law-like connection between the property being measured and the property against which it is measured. Stipulated elements are reflected in the fact that the property being measured is not directly observable.⁴ Ideas of mass, temperature and pressure are not abstract, they are symbolic (theoretical), they are symbols which acquire their meaning owing to the theory they are part of. Since the notion of temperature, for instance, is such that we can never be certain it accurately represents some part of reality, we are free to stipulate any kind of connection between the height of a mercury column and temperature.

⁴ Carlo Giannoni, “Quine, Grunbaum and the Duhemian Thesis”, in: *Can Theories be Refuted*, p.166.

However, this connection is at the same time causal, since it varies owing to the external factors, at least some of which are unknown to us. This brings us to the core of Duhem's conventionalism which can best be summarized in the proposition that there exist laws of nature that we must stipulate as being true.⁵

The second aspect in which Duhem's (and later Quine's) position differs from the classical empiricist one lies in accepting the theory of coherence. Theoretical statements are not to be taken as isolated, but as interconnected inside the web of laws which makes up the complete mathematical representation of experience. Therefore, their truth value cannot be determined independently from the rest of the system. In the holistic picture of science which both Duhem and Quine accept, the units of empirical meaning are not terms or particular statements, but rather the whole theory itself. Quine subscribed to both aspects of Duhem's empiricism. Speaking of scientific theories, he denounced realistic assertions about reality which theories should reflect, and considered theoretical entities as posits which make sense only if seen inside the frame of a theory which postulates them. On the other hand, Quine also insists that our statements about the external world are measured against our sense experience always as a body, never taken in isolation.⁶

Duhem's main goal was to show in which way the procedure of the crucial experiment (which should refute one hypothesis and confirm the other) requires two conditions to be satisfied: (1) there should exist a determined refutation procedure and (2) scientific inference should be susceptible to *reductio ad absurdum* type of arguments. He points out that neither one of these conditions can be met.⁷ Namely, every prediction we make is based not on one, but on many hypotheses in a way that we can never test a single, isolated hypothesis. And even if we could refute an isolated hypothesis (for example, that the nature of light is corpuscular), by doing so we wouldn't be proving true any alternative hypothesis (wave theory of light, for instance). The only thing arrived at by refuting the hypothesis H is its negation $\neg H$, the last being not a single hypothesis but the potentially infinite disjunction of hypotheses.

The reason for this is as follows. By conducting an experiment, the physicist implicitly subscribes to a number of theories without which the experiment itself could not be conducted. In the case the prediction he made fails to come about, the only thing he can do is to claim that some of the statements he based his prediction on were wrong,

⁵ *Ibid.*, p.167.

⁶ *Ibid.*, p.166.

⁷ Larry Laudan, "Grunbaum on the Duhemian Argument", in: *Can Theories be Refuted*, pp. 155-156.

without knowing which one it is. No crucial experiment can decide between two and only two hypothesis.⁸

Let H be the hypothesis which we need to test. By itself, H implies no observational statements. To infer its observational consequences, we must conjoin some initial conditions A to it. Aside from that, since the hypothesis H rarely represents an isolated system by itself, we must also assume some background theory T . Let us assume for the moment that $H \wedge A \wedge T$ implies the observational statement O . Assume also that, after the crucial experiment has been conducted, we observe that $\neg O$ is the case. Does this provide justification for the claims that the hypothesis H is false?

By no means so. We can only conclude that there exist at least one false claim amongst the $H \wedge A \wedge T$, and there is no *a priori* reason that this must be H itself.⁹ This made Duhem conclude that “physical science is a system which must be taken as a whole”¹⁰ so that the “comparison is established necessarily between the whole of theory and the whole of experimental facts”.¹¹

For example, we could raise the question¹² of the structure of physical space. Is it Euclidean or non-Euclidean? In the three dimensional space we can differentiate between three classes of geometries with constant curvature. Loosely speaking, all three classes are based on the first four postulates of Euclid, but each provides its own parallel postulate. The geometry with zero curvature is the usual Euclidean (or parabolic) geometry, while the term “non-Euclidean” is used to refer to both hyperbolic and elliptic geometries.¹³ As already said, the crucial difference between Euclidean and non-

⁸ Duhem, Pierre *The Aim and Structure of Physical Theory*, p. 258.

⁹ In some cases we could even consider taking $\neg O$ as false. We could always, in principle, denounce our experimental results although this should not be the practice of scientists.

¹⁰ Duhem, *op.cit.*, p.187.

¹¹ Duhem, *ibid.*, str. 208.

¹² There are many questions of this kind. For example, is the structure of time linear or cyclic, or how do we interpret quantum phenomena. Regarding the later, it is claimed that different interpretations of the formalism of quantum mechanics give rise to empirically equivalent but different theories, which explain the world in terms of different principles and mechanisms. Perhaps the most striking difference lies between the Copenhagen interpretation, according to which particles cannot have the exact position and momentum at the same moment, and the Bohmian interpretation, according to which particles always have the exact position and speed, but these cannot be simultaneously measured. Namely, quantum phenomena which are susceptible to experimental observation can be consistently and accurately explained on the basis of more than one mathematical formalism. The orthodox quantum formalism and the Bohmian interpretation do share some crucial common assumptions: Schrödinger's equation and Born's rule (the quantum equilibrium hypothesis). This makes their empirical content (empirically observable consequences) equivalent and is a basis for observational indeterminacy. See: Darrin W. Belousek, "Underdetermination, Realism, and Theory Appraisal: An Epistemological Reflection on Quantum Mechanics", *Foundations of Physics*, vol. 35, n° 4, 2005 p. 670.

¹³ The hyperbolic geometry was created by the Russian mathematician Nikolai Lobachevski, although it is necessary to point out that the Hungarian mathematician Janos Bolyai and Carl Friedrich Gauss have arrived

Euclidean geometries lies in the fifth postulate of Euclid. In hyperbolic geometry this postulate is replaced by the following one:

*For any given line l and point P not on l , there are at least two lines through P that do not intersect l .*¹⁴

In elliptic geometry, however, Euclid's postulate is replaced by the following one:

For any given line l and point P not on l , there exist no line through P that does not intersect l .

However, to ensure the consistency of the system in question, after adding this axiom some minor alterations must be made to other axioms as well.

Wherein lies the problem? We shall illustrate this with an example proposed by Poincaré. He made the following thought experiment: imagine the world as located inside the circle c , in which the speed of light in each point is inversely proportional to its distance from the circumference of c . Inside this imaginary world light rays will have the form of circular arcs which are perpendicular to the circumference of c . This apparently shows that the world in question is governed by the laws of hyperbolic geometry. However, instead of viewing light rays as non-Euclidean lines, we could also picture them as Euclidean circles perpendicular to c . The resulting geometry, then, is Euclidean. Therefore, one and the same physical situation can be described using different geometries, given that physical entities (light rays, in our example) are correlated with different notions of geometries we investigate.¹⁵

In the same vain as Poincaré, Einstein also thought that two different geometries can make up systems which are in operational concordance with experience. Since geometry and physics are closely connected, neither one is susceptible to empirical testing taken in isolation. Geometry tells us nothing about the relations between real entities, and can only do so if physical laws are added to it. However, physical theories are not derived

at the same results at the approximately same time. The discovery of elliptic geometry is the product of work of German mathematician Bernhard Riemann, the student of Gauss. The consistency of hyperbolic geometry was proved by the Italian mathematician Beltrami in 1868. He showed that if the Euclidean geometry is consistent, then the hyperbolic geometry is also consistent. Since no one doubted the consistency of Euclidean geometry for more than two thousand years, this was a pretty efficient way to shelter these new, "strange" geometries from further attacks.

¹⁴ Both line l and point P lie in the same two-dimensional plane, of course. The same holds for the axiom of elliptic geometry which follows.

¹⁵ See Henri Poincaré, *Science and Hypothesis*, New York, The Walter Scott Publishing Co. Ltd, 1905, pp. 42-83.

from experience: their statements are, more often than not, freely chosen and are only later brought into operational concordance with experience through appropriate modifications. According to Einstein's view, the question about the geometry of the physical world can be answered if we have a *determined* rigid body. However, for the determination of the rigid body we must have a *determined* geometry. In this way we enter a kind of vicious circle: measuring rod, as a rigid body, cannot even be defined without assuming a physical geometry of the world, since geometry is by necessity needed for calculating corrections without which measuring is not even possible. In this way, the question about the real geometry of the world becomes an empirically meaningless question.

We could represent this situation in a different manner and show that refutation is possible in some cases, although this would not imply that discarded theories are forever abandoned. For example, we can notice that in 1543, empirical evidence was not sufficiently strong to support Copernicus' over Ptolemaic system. Empirical discoveries later made by Brahe and Galileo could not be reconciled with the latter, and by discovering the laws of planetary motion Kepler further improved upon Copernicus' theory. The Copernican theory was crowned by Newton, who showed that inertia and gravity account for elliptic paths of planetary motion.¹⁶ What this shows is that elimination is possible, since the theories have different kinds of limitations. Available facts are just one of those, apart from coherence, plausibility of explanation etc.

In a situation like this, the most that Duhem would be able to claim is that there is no decidable refutation procedure present. Namely, his aim was not to show the impossibility of refutation, but the undecidability of refutation which he saw as the consequence of holism – the claim that physicist can never test a hypothesis taken in isolation from the theoretical framework he is working in. Duhem is not speaking about the possibility of constructing an alternative for every single theory, since this is not part of the practice of scientists. In case we have two hypotheses at hand (for Duhem, these are most usually the old and the new theory) which explain the results of an experiment equally well, and the logic is silent, common sense will be the judge as to which hypothesis to accept. The undecidability of theory choice is, therefore, always the temporary trait of scientific practice. The questions of ontology or the ontological

¹⁶ See: F. Weinert, *Copernicus, Darwin, & Freud: Revolutions in the History and Philosophy of Science*, Oxford, Wiley-Blackwell, p. 68.

commitments of theoretical language are not the subject of Duhem's investigations. Explicitly, he discards them as barren and passing work.

In view of all this, we can distinguish between two forms of Duhem's thesis:

- (1) The logic of refutation and confirmation of an empirical hypothesis H must apply to a web of hypotheses of which H is a part.
- (2) No hypothesis H can be isolated from this web of hypotheses and additional assumptions for the sake of confirmation and refutation.

In his critique of Duhem's thesis, Grünbaum points out two things. Namely, Duhem's first claim expresses elementary logical truth that if the observational statement O is implied by the conjunction of hypothesis H and additional assumptions A , then its falsity cannot be a ground to conclude $\neg H$, but only $\neg(H \wedge A)$. Refutability of H is undecidable in the following sense: the falsity of H does not follow deductively from the premise $((H \wedge A) \rightarrow O) \wedge \neg O$. Duhem's second claim is, according to Grünbaum, more problematic. Observational statement $\neg O$ which is incompatible with O , enables us to claim that H is true and that A is false, since it allows for a theoretician to keep H and modify A , so that both H and A' , a modified version of A , together imply $\neg O$. The first claim, although correct, is not at all sufficient to show that the falsifiability of H is unavoidably inconclusive.¹⁷ It cannot rule out the possibility of some nontrivial¹⁸ variant A' which will, together with H , imply $\neg O$.

In view of all this, Grünbaum concludes that this thesis is untenable in its nontrivial form, i.e. that the thesis of non-susceptibility of hypothesis to refutation is neither logically nor scientifically justified. Contrary to that, Grünbaum claims that conclusively refuting experiments are possible, and that crucial experiments are a fact in the history of physics.

The basic question here is why do we think that there always exists an A' with the help of which we could save the hypothesis? Namely, if we are unable to prove that there always exists some nontrivial A' for each H and $\neg O$, then we need not believe that every refutation procedure is undecidable. Duhem never claimed that $\forall H \forall O \exists A' ((H \wedge A') \rightarrow O)$ holds. His position was far more moderate. He did not claim that we can save each and every hypothesis, but that if it's not proved that the hypothesis cannot be saved, then

¹⁷ See Adolf Grünbaum, "The Duhemian Argument", in: *Can Theories be Refuted*, pp.116-117.

¹⁸ A special problem arises when we consider when we are to take A' as a nontrivial modification. Without going into subtle considerations, for the purposes of this work it will be enough to label a modification nontrivial if it is not purely based on language variations.

we cannot refute it either.¹⁹ The burden of proof here lies with those who deny H , and they need to show that there does not exist an A' which would make H compatible with $\neg O$. Accordingly, Laudan differentiates between two forms of **DQT**: a stronger one, which is the basis of Grünbaum's attacks, and a weaker one which he argues was held by Duhem.

The stronger thesis claims the following: For every hypothesis H and every observational statement O , there exist a set of nontrivial auxiliary assumptions A' , such that H and A' entail $\neg O$.

The weaker thesis, on the other hand, claims: In the absence of a proof that no appropriate hypothesis-saver exists, then $\neg O$ is not a refutation of H , even if $(H \wedge A') \rightarrow O$.²⁰

The stronger thesis Grünbaum criticizes is not Duhem's, but Quine's. Quine held that theory could always be reconciled with recalcitrant experience by making appropriate changes in the rest of the system.²¹ In this context that would mean that there always exists a group of theoretical assumptions A' . Duhem, however, does not hold that A' always exists. His thesis is weaker and it claims only that a physicist can never be certain that no such A' exists.²²

Grünbaum's aim was to show that **DT** is correct only in a trivial sense in which drastic changes in other parts of the system can be made only through an *ad hoc* change of language rules. On this view, the necessary condition of non-triviality of **DT** is that theoretical language be semantically stable in all the relevant aspects. Grünbaum thought that we need some kind of a theory of change and persistence of meaning inside the web. It needs to be pointed out once more that the thesis which Grünbaum ascribes to Duhem is really Quine's.

Quine's further modification of **DT** makes it more detached from the above interpretation. Most notably, Quine's position is characterized by a more radical form of holism. As we already noted, Duhem held that physical theory makes up an interconnected body and that we cannot test hypotheses taken in isolation from it. To this body Quine further adds both logic and mathematics, which means that when we test certain hypotheses this test includes all the elements this body contains. The unit of

¹⁹ Larry Laudan, "Grünbaum on the Duhemian Argument", p.158.

²⁰ *Ibid.*, p. 159.

²¹ W. V. Quine, "Two Dogmas of Empiricism", p. 51.

²² Gary Wedeking, Duhem, "Quine and Grünbaum on Falsification", in: *Can Theories be Refuted*, pp. 178-179.

empirical significance is, to use Quine's words, the whole of science.²³ In contrast to Quine's thesis, **DT** has a historical character, it gains ground from the history of science. Quine works only inside the framework of contemporary science, most notably contemporary physics. Duhem's thesis also has a limited scope, for it does not concern psychology, for example, while Quine's thesis concerns *the whole body of knowledge*.²⁴ Apart from that, **DT** incorporates a delimitation between science which is formulated in a mathematical language and natural language that is not subsumed by it, while **QT** concerns the natural languages as well. As is well known, Quine is questioning the difference between synthetic and analytic propositions, and according to his view mathematics (which on Duhem's view is analytic and not susceptible to verification and hence refutation) is just one among many myths we created which, unlike Homer's gods, has shown greater pragmatic value.

Contrary to Duhem, Quine does not subscribe to a position which acknowledges that we cannot know which statements of our theory should be cast out in the light of refuting experience – he takes this even further by claiming that we are presented with a choice concerning the nature of the revision of our theory in this kind of situation for “any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system...and conversely...no statement is immune to revision”²⁵

We can point out some of the difficulties which this interpretation faces by using two simple examples. Let *T* be Ptolemaist theory and let *A* be a statement of existence of the unchangeable sub-lunar sphere of Aristotelian cosmology. Also, let *H* be a statement that there are no stars that do not revolve around the Earth (i.e. that Jupiter has no shepherd moons) and let $\neg O$ stand for Galileo's discovery of the moons of Jupiter. This last discovery, together with the discovery of phases of Venus and Brahe's observation of supernova in 1572 are showing the mutability of the sky. These discoveries are hard to incorporate inside a geocentric model, and it is hard to see how we would be able to change the background assumptions in order to explain the empirical results. If we were to replace *A* with *A'* stating the mutability of the sky, this would collapse the geocentric model itself. We could deny Galileo's discovery to the point that empirical evidence can

²³W. V. Quine, “Two Dogmas of Empiricism”, p. 50.

²⁴ Vuillemin, “On Duhem's and Quine's Thesis” in: *The Philosophy of W. V. Quine*, La Salle, Open Court, 1986, p. 599.

²⁵ W. V. Quine, “Two Dogmas of Empiricism”, p. 51.

be ignored. However, once empirical evidence is credible enough, this kind of strategy becomes dogmatic.

On the other hand, sometimes it is easy to shed additional assumptions. Let T be Copernican astronomy, A an assumption on uniform, circular paths of the sub-lunar sphere, H a hypothesis on circular motion of planets around the central sun and $\neg O$ a statement stating Kepler's discovery of elliptical, non-uniform path of planetary motion. Kepler rejected A replacing it with A' , stating the existence of non-uniform motion, since he considered Copernicus' theory essentially correct. If he would to reject Copernicus' theory, that would mean going back to some of the ancestral theories which were not compatible with observational data. Therefore, he changed the background assumptions so that they, together with T , imply $\neg O$. However, it is questionable whether this is a case of saving a theory at any cost, since this theory was, in his view, essentially correct.

What this shows is that the situation is not as simple as the simplified scheme Quine suggested implies. These, however, are not the only problems it faces. For example, Lakatos has pointed out that if we can renounce any statement facing a refuting experience, then we are putting in question any real progress in science, which by itself collapses into conventionalism. Quine does not, however, suggest that theory should be adjusted in an arbitrary manner (the possibility of denouncing the refuting experience as illusory included) to "save the phenomena". He simply claims that even the so called analytical statements (those of mathematics and logic included) are as susceptible to revision as the empirical statements, although to a different degree. There were statements to the effect that the law of excluded middle of classical logic be rejected in order to better describe the phenomena of quantum mechanics, Euclidean geometry was replaced with Riemannian in Einstein's theory and so on. If these were valid moves, than logical and mathematical statements are not as sacrosanct as it is usually thought, and future revisions in this field are very much possible. Moreover, Quine also suggests a criterion that should serve as a guidance for revision: they should be brought about in such a way that "the whole of the system is disturbed as little as possible". Although it could be argued that this criterion is not completely satisfying (for example, it could be seen as too conservative for it obviously favors existing established theories), it does possess an advantage that it rules out those interpretations of Quine's philosophy that are ascribing to him the view that established theory must be defended at any cost.

Despite everything that has been said, it seems that Quine's modification of Duhem's thesis **DT** is relatively free of problems. It is often pointed out that it expresses a simple logical fact applied to science. Namely, *modus tollens* shows only that the antecedent of an implication must be false, not which part of it is responsible for this happening. If we are to agree with Quine in denouncing the difference between analytical and synthetic propositions, which implies that mathematical and logical propositions have the same status as their empirical counterparts, his first modification of **DT** also seems acceptable. However, it is certainly stronger and has a broader scope than **DT**.

In his later works Quine formulates his thesis as a thesis of *underdetermination* and points out that it concerns the relationship between scientific theories and evidence upon which those are based. Roughly speaking, it claims that there can exist numerous opposing theories explaining the same body of evidence. If we are presented with the same evidence, it is possible to construct two different theories which will have the same explanatory power.

A simple proof of the thesis of underdetermination can be constructed in the following way: if we are given a hypothesis (H), a set of statements about the evidence (P) and a certain number of statements we can add to these as an additional hypotheses ($S_1, S_2, S_3 \dots$), then the following holds:

1. $H \rightarrow P$
2. $H \wedge S_1$
3. H
4. P
5. $(H \wedge S_1) \rightarrow P$

What we have shown is that if a theory implies an evidence, then we are able to construct a whole sequence of theories with the same effect. In other words, there are infinitely many theories ($H \wedge S_1, H \wedge S_2, H \wedge S_3, \dots$) which imply the same evidential statements. This is, in a nutshell, what Quine understands by underdetermination of physical theories. However, it should be kept in mind that what we have just outlined is merely the simplest form of underdetermination.

This modification of **DT** made by Quine (which is also called the "strong interpretation" of **DT**, and which we will refer to as the thesis of underdetermination in what follows) was the subject of much criticism. What is in play here is the *equality thesis* which claims the following: every theory is equally well confirmed by evidence as any

of its rivals. In other words, any theory can be reconciled with resisting evidence by making suitable adjustments in other assumptions we hold about the nature.

As already pointed out, if theories imply observational consequences only with the help of additional assumptions, then the theory together with suitable additional assumptions can always be reconciled with the resisting evidence. The consequence of this is that for any evidence and any two rival theories, T and T' , there are suitable additional assumptions A such that $T' \wedge A$ will be empirically equivalent to T (together with its additional assumptions). Therefore, evidence cannot decide between them.

Against differentiating between Duhem's and Quine's thesis, we could point out the fact that Quine relied on **DT** as an assumption in his derivation of the underdetermination thesis. Namely, he assumed that everything implied by **DT** is also implied by the thesis of underdetermination. As already pointed out, in "Two Dogmas of Empiricism" and "Naturalistic Epistemology" Quine insists that the whole field of science is underdetermined by the side conditions of experience in such a way that we are always presented with a choice regarding which statement we want to change in the light of disconfirming evidence. On this point, Duhem's holism is made equal with underdetermination.

Contrary to this, we find these two to be non-equivalent, since the thesis of underdetermination does not concern intra-theoretical dependencies of terms and sentences, as **DT** and **QT** do. **DT**, as well as the weaker interpretation **QT**, claims that the conjunction of laws implies observational consequences, while underdetermination can be seen as a claim that evidence implies different theories. Therefore, these two theses are true for any theory, even if there is no more than one such. For the case of underdetermination to occur two conditions must be satisfied by necessity: (a) at least two empirically equivalent theories must exist and (b) those theories must be logically incompatible and cannot be made logically compatible by mere reconstruction of their theoretical predicates. Both Duhem's and Quine's thesis are consistent with the thesis of underdetermination being either true or false. They are concerned with truth conditions of singular theoretical statements. In other words, even if underdetermination is not present, they do not cease to hold.

We can therefore conclude that there are differences between these two theses, and that Quine modifies **DT** to arrive at his thesis of underdetermination. The upshot of Quine's arguments is not that **DT** implies underdetermination by itself. To arrive at the

case of underdetermination we need additional assumptions: we must have rival, empirically equivalent and logically incompatible theories.

QT is the stronger thesis, because it is not limited to physical theories and also because of Quine's rejection of the analytic/synthetic distinction. Quine's thesis incorporates what we have called **DT** at the beginning, but since **DT** claims nothing about the truth of mathematics and logic, its scope is wider. They are not identical since **DT** claims nothing about the changes that must be made inside the system to save some singular statements. **DT** is just the thesis of holism which claims that we cannot verify isolated parts of the system, and since the system is such that apart from theory also includes additional background assumptions, we cannot localize the source of error in the case of a failing prediction. **QT** is, in its weaker form, closer to **DT** but is still stronger than it. In its stronger form, in which we can understand it as a thesis of underdetermination, it claims something more than what Duhem wrote about the crucial experiment.

Duhem showed why crucial experiments in science are not possible: as a consequence of holism we cannot speak about decisive procedure of refutation and confirmation. When we have two rival theories, even if we would be able to refute one of them, we would not be proving the other one true by that same act. Quine's thesis of underdetermination has a different content. It brings about the possibility of constructing alternative, incompatible theories, and so can be understood as an algorithm for creating empirically equivalent theories. Although Quine moved away from this form of thesis of underdetermination in his later works, it should be kept in mind that even the weaker forms of underdetermination he came to find acceptable are not equivalent with Duhem's or Quine's holistic thesis. Namely, these last two hold even if underdetermination is not present. On the other hand, underdetermination is a phenomenon that is not a consequence of holism alone, but more importantly of the imperfection of human conceptualization of reality, of the limitations of models we apply or, in short, of the fact that our theories always go above and beyond what is given in empirical evidence. Therefore, although holism might be a necessary condition of underdetermination, it is certainly not its sufficient condition.

Pierre Duhem's Consideration of Infinity in the Middle Ages*

Science and the history of science

The great French thinker Pierre Duhem became a scientist in the nineteenth century, but later turned to the philosophy of science and devoted himself exclusively to it in the twentieth century. In his first two books (*Le mixte et la combinaison chimique* from 1902, and *L'évolution de la mécanique* from 1903), dedicated to the history of science, Duhem does not pay much attention to medieval thought. This was a common viewpoint at the time. For many scientists and historians of science, even at the beginning of the twentieth century, the term 'medieval science' was utterly nonsensical, almost like the 'round square'. The Middle Ages were thought to be a time of theology and metaphysics, while science was thought to appear in the Early Modern Era. Obviously, a time in which there was no science could not be of interest for the history of science.

Jan Łukasiewicz also held this position. He claimed that the history of logic available in his time was badly written. The reason for this lies in the fact that available histories of logic were not written by logicians, and only a good logician can identify what is logically significant in documents preserved from the past.¹ If this view is correct, then Duhem was in a good position since he was first a scientist, then a philosopher of science, and only towards the end of his career became a historian of science. This, however, did not mean that he was immediately able to recognize the important elements of medieval thought which gave way to modern science. Indeed, Duhem followed the accepted view that there was no science in the Middle Ages, and in his first two books dedicated to the history of science did not discuss any of the scientific issues addressed by medieval thinkers.

* This paper is the result of my work within the projects *Dynamic Systems in nature and society: philosophical and empirical aspects* (number 179041) and *Logico-Epistemological Foundations of Science and Metaphysics* (number 179067), financed by the Ministry of Education, science and technological development of the Republic of Serbia.

¹ Jan Łukasiewicz, 'On the History of the Logic of Propositions', in: Jan Łukasiewicz, *Selected Works*, Amsterdam – London, North-Holland Publishing Company, 1970, p. 198.

The turning point in this respect was the book *Les origines de la Statique* (1904), in which he especially discussed medieval conceptions and, according to some historians of ideas, thereby oriented the attention of other researchers towards the science of the Middle Ages. Between 1906 and 1913, Duhem wrote and published a series of essays on the impact that thinkers from the Middle Ages had had on the works of Leonardo da Vinci.² From 1909 to 1916 (the year of his death), Duhem wrote his most significant work, *Le Système du monde*.³ He envisioned a grandiose text in twelve-volumes, which would present the history of cosmology from Plato to Copernicus. This endeavour was cut short by his death, though during his lifetime he managed to finish nine volumes. The first five were published between 1914 and 1919, and the further four volumes were published in the 1950s.⁴

The destiny of Duhem's discussion of infinity in the Middle Ages

In the last two works we have mentioned, Duhem advocated the thesis of continuity between late medieval and early modern science. This thesis was considered unacceptable in his time, and it is still challenged today. If this theory is difficult to defend when it comes to other issues, Duhem has quite convincingly shown that this is not the case when it comes to the problem of infinity in the late Middle Ages, especially in the first half of the fourteenth century. Historians of mathematics have rarely taken a look at the problem of actual infinity in the scholastics, whether before or after Duhem. Most historians argue that the problem was set in ancient Greece, and that the late scholastics only repeat Aristotelian teaching on the matter. Yet, we know that historians of mathematics often mention Galileo, and then look back to the second half of the nineteenth century, to the teachings of Bolzano, Dedekind and Cantor.

It seems that some responsibility for this state of affairs lies with Bolzano and Cantor. In his book *Paradoxien des Unendlichen*, for example, Bolzano says that the relations between infinite sets, as well as the relations between these sets and their proper subsets, were not only neglected in the Middle Ages — they were forgotten in his own

² Pierre Duhem, *Études sur Léonard de Vinci*. 3 volumes, Paris, Hermann, 1906-1913.

³ It is an incredible fact that he wrote the three-volume book *Traite d'énergetique* during the same period.

⁴ Pierre Duhem, *Le système du monde, histoire des doctrines cosmologiques de Platon à Copernic*, 10 volumes, Paris, Hermann, 1913-1959.

time too.⁵ When it comes to the former, Cantor says that “as it is well known, throughout the Middle Ages ‘infinitum actu non datur’ was treated in all the scholastics as an incontrovertible proposition taken from Aristotle.”⁶ It appears that the contribution these authors made to the debate on actual or completed infinity, and to the problem of infinity in general, was a signal to the historians of mathematics. It seems that historians of ideas, precisely because of this, did not search for contributions to the theory of infinity in the texts of the Middle Ages.

Such view of the historical significance of the way in which Late Scholastic scholars considered the issue of actual infinity, was challenged by Duhem who saw in Gregory of Rimini the central figure of this change. However, many historians of mathematics, who have dealt with the same subject, rather unexpectedly do not take into account Duhem’s results.

At first glance, we might fail to recognize the importance of one historian of the Middle Ages: Annelise Maier. In her paper, ‘Discussions about the actual infinity in the first half of the 14th century’, she also fails to mention Duhem’s extensive work on the same problem in *Le Système du monde*. However she does refer to the work he dedicated to Leonardo da Vinci, and there are clear reasons for that. First, the importance the late Scholastics gave to questions of actual infinity had been noted by Duhem in his essays on da Vinci, but volume VII of *Le Système du monde*, in which Duhem discusses in detail the issue of the actual infinity, was published in the 1950s. Maier’s paper, however, was published in 1947,⁷ and Duhem’s book was not available to her.

Yet some historians of ideas believed that Maier was the first historian of science to stress the importance of medieval discussions of actual infinity. This implicitly follows from a number of papers by John Murdoch, who, in his extensive opus on this topic, mentions Annelise Maier, but nowhere speaks of Duhem. In his study of Gregory of Rimini, John Thakkar, for example, does not mention either Duhem or Maier. He simply says that ‘John Murdoch first brought to light the remarks on infinity made by the Augustinian hermit Gregory of Rimini’.⁸

⁵ Bernard Bolzano, *The Mathematical Works of Bernard Bolzano*, Oxford, Oxford University Press, 2004, p. 16.

⁶ Georg Cantor, *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*, Berlin, Springer, 1932, pp. 173-174.

⁷ A. Maier, “Diskussionen über das aktuell Unendliche in der erste Hälfte des 14. Jahrhunderts”. First published in *Divine Thomas*, Serie III, Bd. 24, 1947. Now available in: A. Maier, *Ausgehendes Mittelalter*, Roma, Storia e letteratura, 1964, pp. 41-85.

⁸ J. Thakkar, “Infinity, Continuity and Quantification: Gregory of Rimini on the Crest of the Fourteenth Century”, 2004, <http://users.ox.ac.uk/~ball2227/files/continuity.pdf>, p. 1.

Duhem's consideration of infinity in the late scholastics

As we have mentioned, Duhem's work on medieval debates about infinity was published in the seventh volume of *Le Système du monde*, all relevant parts of which can now be found in English translation, published in 1985.⁹ This paper will refer to the English translation.

Although Duhem's presentation begins, as one would expect, with Aristotle, the principal figure around which the story revolves is Gregory of Rimini (1300-1358),¹⁰ and his understanding of the actual infinity. When it comes to the fourteenth century, Duhem claims that essays of most philosophers who preceded Gregory of Rimini "are not really doctrines". In other words, they are just "attempts" that gained historical significance from the fact that they 'did rise to a theory in which the power of the fourteenth century logic can be felt: Gregory of Rimini's theory." As such, they seem more interesting "because they appear as the beginning of the doctrine which the Augustinian master completed".¹¹ Duhem sees the way that Rimini addressed the issue of actual infinity as important because it is a step forward on the historical road towards a Cantor-style theory of infinite sets, and that is how his teachings gained historical significance. Having this in mind, one can understand better his point that 'it is interesting to read the first few pages of Georg Cantor's *Theory of Transfinite Numbers*. There is a clear affinity between the thoughts of the two powerful logicians ... Gregory of Rimini certainly glimpsed the possibility of the system Cantor constructed."¹² To demonstrate this claim, Duhem cites Rimini:

"With respect to infinite multitude, we have used the two words *how much* and *so much* (*quot et tot*); similarly, nothing restrict us from using them with respect to infinite magnitude (*quantum et tantum*). If, for example, one follows the opinion of the Philosopher and one asked how much time has preceded the present instant, one could reply, an infinite time. The infinite can therefore reply to the

⁹ Pierre Duhem, *Medieval Cosmology – Theories of Infinity, Place, Time, Void, and the Plurality of Worlds*, ed. and transl. R. Ariew, Chicago, The University of Chicago Press, 1985.

¹⁰ Gregory of Rimini was on two important points connected with France. He first studied theology at Paris from 1322 until 1329, then he returned to Paris in late 1340 or in 1341 in order to lecture on the *Sentences*.

¹¹ Duhem, *Medieval Cosmology*, p. 109.

¹² *Ibid.*, p. 112.

question, *how much (quantum)?*, and it is a quantity (*quantum*), if, as one asserts, everything that answers the question, *how much?*, is a quantity.

But perhaps one wishes to use the words *how much* “quantum” and *so much (tantum)* only with respect to magnitudes that are finite in some measure. In that case, I would say that an infinite magnitude is not a quantity (*quantitas*), but that it is a magnitude (*magnitudo*); similarly, an infinite multitude would not be a quantity, but it would be a multitude. The word quantity (*quantitas*) would no longer designate the most general genus of the second predicament, we would have to construct a new name for this predicament. But this restricted meaning of the term *quantity* is neither useful nor timely.

I therefore state that infinite magnitude is certainly contained in some species of quantity. Consequently, magnitude is divisible into infinite and finite magnitudes; finite magnitude is then subdivided into magnitudes of two cubits, three cubits etc.”¹³

By quoting Rimini, Duhem wishes to draw our attention to the fact that Rimini clearly points at the difference between the mathematics of finite magnitudes and the mathematics of infinite magnitudes. A few pages earlier, Duhem argues that Rimini’s infinite magnitude in fact refers to *transfinite* magnitude.

Duhem displays the long path taken by the medieval thought, and shows how it was influenced by Rimini’s argument. He first presents a set of facts that shaped the discussion of the problem of infinity in the late scholastics. On the one hand, it was the translation of Aristotle’s work, especially of his *Physics*, as well as the comments on Aristotle made by Islamic medieval thinkers. On the other, it was the Christian dogma as articulated in the *Sentences* of Peter Lombard. Most discussions of the problem of infinity were thus based on interpretations of those texts.

Late thirteenth and early fourteenth century thinkers preferred the Second book of the *Sentences*, in which Lombard considers issues that do not belong to theology in the proper sense, such as motion, perception, cosmology, and so on. Rimini discusses this issue in his lectures on the *Sentences*, which he held in Paris in 1334. Sometimes, talking about the *Sentences* was just an excuse to engage in this debate. Rimini’s excuse was John Damascus’ question, which was quoted in the *Sentences*: “in what place were angels created?”, and whether “an angel is in an indivisible or a divisible place?”. But straight after that, he moves on to issues that really interest him: “The first is whether a magnitude is composed of indivisibles ... The second whether is anything indivisible in a

¹³ *Ibid.*, p. 113.

magnitude'.¹⁴ Discussing the issue of God's omnipotence, Lombard raises traditional questions like whether God can sin, die or lie, and so on. But Rimini is really interested in the question of 'whether God, through His infinite power, can produce an actually infinite effect.'¹⁵

In addition to giving a basic literary background, Duhem emphasizes a revised conceptual network that enabled progress in addressing the problem of the infinity. Instead of Aristotle's distinction between *actual* and *potential* infinity, thinkers of his time insisted on a difference between the *syncategorematic* and the *categorematic* infinite, between the infinities treated *distributively* and treated *collectively*, or between the infinite *in facto esse* and infinite *in fieri*. The question of God's omnipotence had an important role in those discussions, and theologians of the time held fast to Aquinas' definition of a God who can do everything that is logically possible.

Duhem also reminds us that discussions of infinity occurred within discussions about other broader issues, such as, for example, the old theme of the eternity of the world. In fact, the antagonism that emerged over the issue is, he says, the antagonism between the two theorems of Aristotelian philosophy, namely:

1. The universe has existed or could have existed from all eternity.
2. Something infinitely large is not possible in actuality.¹⁶

He goes on to say that this antagonism was recognized by the Arabian *Kalam* philosophers, who 'rejected the first proposition because of the second'. It might seem strange that Duhem does not mention John Philoponus (490-570), who discovered this argument, and al-Kindi (c. 801-866) who was the first, in Islamic and Jewish tradition, to affirm this type of argumentation. It could, however, be explained by the fact that Philoponus' consideration of this issue was not known to the thinkers of the time, and that al-Kindi's texts dealing with these questions were similarly unknown, even to Duhem. Al-Kindi's manuscripts were found after Duhem's death, by the middle of the last century, in the library of Hagia Sophia in Istanbul. Duhem, however, draws our attention to what was relatively new in his time. Besides the authors who repeated the *Kalam* argument, which from the impossibility of the actual infinite concludes the impossibility of the

¹⁴ Gregory of Rimini, *Lectura super primum et secundum sententiarum*, ed. A. Damasus Trapp OSA, Venício Marcolino et al., 7 vols, Berlin, Walter de Gruyter, vol. IV, 1979-87, pp. 277-278.

¹⁵ Gregory of Rimini, *op. cit.*, vol. III, p. 438.

¹⁶ Duhem, *Medieval Cosmology*, p. 80.

eternity of the world, as does Richard of Middleton, he also refers to authors such as John Bassols, who from the possibility of the eternity of the world deduces the possibility of actual infinity.

Let's return to Rimini. It does not hurt to remind ourselves again that Rimini's view of the problem of the actual infinite hinges on the example of continuous magnitude or, more precisely, on the possibility of its infinite divisibility.

A new conceptual network, in particular the distinction between the *syncategorematic* and the *categorematic*, which in the Latin context was introduced by Priscian, and in the context of the discussion of the infinite was reaffirmed by Peter of Spain,¹⁷ enabled Rimini to see that infinite magnitude is not something that could be obtained from finite magnitude, but instead is something fundamentally different: as such, it is something transfinite. To illustrate this, Duhem writes that Rimini claims that the term 'infinite' can be used *syncategorematically* and *categorematically*. 'In the first sense', Rimini says, 'the term is equivalent with the phrase "not so great that it cannot be greater", in the continuous case; or "not so many that it cannot be more numerous", in the discrete case. I believe, however, that the proper phrase would be "greater than any finite thing, however great" or "more numerous than any finite number". And categorically, if it is used for continuous magnitudes, "the infinite" is explained by the phrase "so great that it cannot be greater" ... and in the discrete case, "so many that [it is not and cannot be] more numerous". A few lines later he gives an example: "suppose that the world will continue forever as it is now. If someone says 'infinite men will be past', then the proposition is true if the word 'infinite' is taken syncategorematically, since for every finite number [of men], more will be past; but taken categorically it is false, since there will never be more than every finite number [of men] past."¹⁸

Among other major precursors of Rimini's thesis that *categorically* infinite magnitude and multitude are possible, Duhem counts John Duns Scotus. He argues that Scotus gave significant contribution to the whole debate. Duhem cites Scotus: "The words, *equal*, *greater*, and *smaller*, are not suitable for large quantities unless finite. In fact, before one can apply the words equal or unequal to a quantity, one has to divide it into a finite quantity and an infinite quantity. The reason for which a quantity is greater

¹⁷ For more on this see N. Kretzmann, "Syncategoremata, exponibilia, sophismata", in: *The Cambridge History of Later Medieval Philosophy*, ed. Kretzmann, N. Kenny, A. Pinborg, Cambridge, Cambridge University Press, 1982.

¹⁸ Gregory of Rimini, *op.cit.*, IV, p. 294.

than another one lies in the fact that it exceeds it; the reason of equality, in the fact that it has the same measure (*commensurari*). Everything indicates that these concern finite magnitudes. One must therefore deny that an infinity can be equal to another infinity; *more* and *less* also designate differences between finite quantities, and not infinite quantities”.¹⁹

An error made by the majority of predecessors was to think that relations that are valid for finite magnitudes can also be applied to infinite magnitudes. Scotus rejects this, but unlike others who deny the possibility of *larger* or *smaller* infinities, it seems that he rejects the possibility of equal infinities too. Scotus thinks that finite magnitudes are commensurable, that they can be *equal*, *greater*, and *smaller*, but that infinite magnitudes are absolutely incommensurable. From this, it can be clearly concluded that infinite magnitudes are, for Scotus, transfinite.

While mapping this out and approaching Rimini’s point of view, Duhem speaks of the conceptions of Francis of Mayronnes, John Bassols, John Bacon of Baconthorpe, Nicholas Bonet, Robert Holkot, and of many others, and concludes that all becomes ‘clearer when one relates it to the teachings of Gregory of Rimini’.²⁰

Duhem’s interpretation of Rimini’s teaching on the categorematic or the actual infinite

Duhem assumes that Rimini tries to answer the question: can the words *greater*, *smaller*, *whole*, and *part* be used legitimately with infinities, and do they then have the same meaning as when they are applied to finite magnitudes? However, he first examines the two ways one *can* use these words. According to Rimini, they can be used in their common meaning and in their proper meaning. Duhem summarizes Rimini’s understanding by saying that “according to its common meaning, the *whole* is a *part* and anything else not contained in the part; according to its proper meaning, the *whole* is a *part* and a number of things of determined magnitude.”²¹ He clearly notes the significance of the distinction which Rimini makes between “anything” and

¹⁹ Duhem, *Medieval Cosmology*, pp. 89-90. Duhem quotes according to Joannis Duns Scoti *Scriptum Oxoniensis*, lib. II, dist.I, quaest. III.

²⁰ Duhem, *Medieval Cosmology*, p. 101.

²¹ *Ibid.*, p. 110. Gregory of Rimini, *op. cit.*, III, pp. 457-458.

“determinate”. In this regard, he cites Rimini’s clarification which says that “for a thing to be called a *whole* with respect to another”, in the proper sense of the word, “it is not sufficient that it contains the other, as is assumed in the first way [the common meaning], but it has to contain a determined number of things of determined magnitude (*to tanta*) not contained within that which is included”. In the proper sense of the word, writes Rimini, “an infinite multitude cannot be either a *whole* or a *part* of another infinite multitude; there is no determined number of groups of such unites (*tot tanta*) contained in one of the multitudes and not in the other, for each of them contains an infinite number of groups of such unites (*infinites tantum*) or an infinity of such groups in which one can count an infinity of unites (*infinita tanta*).”²²

Rimini says something similar about the meaning of the words *larger* and *smaller*. He also makes a distinction between the use of these words in their proper sense and in their improper sense. In their proper sense, he writes, ‘a multitude is said to be larger than another when it contains not only a number of units as large as the other, but also a larger number (*tantundem et plures*). A multitude is said to be less than another when it contains a lesser number of units (*pauciores*)’. But if these words are used in their improper meaning, then, says further Rimini, “if a multitude contains all the units of another multitude and also some different units than the former, one says that it is larger than the former multiple, even though it does not contain a *larger number* of units (*pluras unitates*) than the other multitude.” But these words cannot in the proper meaning be used “in the comparison of infinities to one another; one must use them only to compare finite magnitudes. One can still say that an infinite is larger than a finite magnitude and that a finite magnitude is smaller than an infinite.”²³

From all that has been said, it follows that infinite magnitudes (or multitudes) cannot be put in relations of size, of being larger or smaller. Does this mean, then, that they are equal? Duhem draws our attention to the fact that Rimini implicitly shows that those traditional examples, which were offered as arguments for the thesis that all efforts to actualize infinity lead to the contradiction of unequal infinities, are not good. Many thinkers in the past, starting with John Philoponus, gave examples which were supposed to violate the axiom that all infinities are equal, and that infinities of different sizes cannot

²² Duhem, *Medieval Cosmology*, p. 111. Gregory of Rimini, *op. cit.*, III, p. 458.

²³ Gregory of Rimini, *op. cit.*, III, p. 458-459.

exist. From the point of view of Rimini and Cantor, however, all infinities are actually equal.

Rimini, however, tried to offer an argument that the *part* of a continuous *whole*, in terms of its infinite divisibility, can be *equal* to that *whole*. In fact, he says that the proper use of words *whole* and *part* shows that “no infinite multitude is a whole or a part with respect to [any other] multitude, since none includes a certain amount without any other including that amount, for they all include infinitely much or an infinite number of things”. Thus, Rimini continues, “for however many pairs or threes the multitude of the whole includes, the multitude of part includes just as many”. Therefore, he concludes, the “multitude of the whole includes no more than that of the part.”²⁴ As such, infinite magnitudes are neither greater nor smaller than each other. Instead, they are exactly equal.

Predecessors have often claimed that *infinity* cannot be actual because it cannot be traversed in a finite time. In his presentation, Duhem draws our attention to the way Rimini solves this problem. He first interprets the traditional view, according to which, he believes, ‘infinity is something such that when one takes any part of it whatever, there always remains another part to be taken’. That means that infinity is something that can never be completed. In the language of the time, infinity can exist *in fieri*, but not *in facto esse*. Rimini thinks that this can be understood only “by admitting that the parts taken successively are all of the same magnitude and that they are all taken in equal times”. It is clear that in this way infinity could never be traversed or, as Rimini says, ‘there will always remain something to be taken of this infinity, and it will never be taken in its totality’. In this case, the traversing of infinity requires infinite time. But if one wants actually to traverse infinity, one has to do it in finite time.

What is Rimini’s solution? It is based on a different organization of relations between parts of time and parts of the infinite: “But once equal parts of the infinity are not taken in equal times, but in times whose durations decrease in geometric progression ... there is no longer any inconsistency in the infinity being taken in its totality”.²⁵ Yet if the final long period of time divides, no matter how small its parts are, the number of parts will always be finite. But in this case, the infinite number of parts cannot be put into one-to-one correspondence with the finite number of parts of time. The only solution,

²⁴ Gregory of Rimini, *op. cit.*, III, p. 458. Unfortunately, Duhem stops quoting Rimini’s discussion regarding those relations just before he makes the main point under consideration. As such, Duhem does not state the section I referred to in the last paragraph.

²⁵ Duhem, *Medieval Cosmology*, p. 116.

Rimini thinks, is to divide a finite period of time into unequal parts. In this way we can, just like Zeno, come to an infinite number of parts of time, whose duration is finite, for example, one hour. However, that division should be an orderly one, for otherwise it could not be put into one-to-one correspondence. The duration of the parts of that finite time should decrease exponentially or, as Rimini says, in geometric progression. This type of division is usually called the division into proportional parts.

One further question remains, one that was common at the time. This is the question of *who* could traverse infinity in a finite period of time? From the above arguments, it could easily be inferred that God might exceed infinity in an hour, for example, taking half an hour for one part of infinity, then a quarter for the second part, an eighth for the third, and so on. At the end of one hour infinity would be exceeded. Rimini provides an example in which he says that God “could have created a stone in each of the proportional parts”. And “since the multitude of these proportional parts is infinite, by the end of the hour there will result an infinite stone.”²⁶ Duhem reminds us that this kind of argument can be objected to, namely, that “if one assumes a continuum divided into proportional parts, one is not allowed to state that one takes *all* the proportional parts of this continuum, parts whose multitude is infinite, for there will be a part taken last”.²⁷ But there is no proportional part which is the smallest or the last.

Duhem states that Rimini thinks that this objection has weight if the point is understood *syncategorematically* or *distributively*. It assumes that the successive parts of a continuum should be taken *in equal times*. But if the infinite divisibility of a continuum cannot be traversed or completed *syncategorematically* or *distributively*, it can be completed *collectively* or *categorematically*. It is impossible to traverse all parts of infinity *distributively*, because, Rimini tells us, “it is impossible that *all the parts* of this infinity can be taken all together; whatever number of parts already taken ... they are still parts of the whole containing them and, consequently, they have a part, or other parts, outside them.” But if we do not take “all the parts” *distributively*, but *simultaneously* or *collectively*, then, writes Rimini, the ‘*set of things* [*omnia*, instead of *omnes partes*] of which each is a part of this whole constitutes the whole; inversely, the whole is *the set of things* of which each is one of its parts. In this proposition, the words *set of things* (*omnia*) are taken collectively”.²⁸

²⁶ *Ibid.*, p. 114.

²⁷ *Ibid.*, p. 116.

²⁸ *Ibid.*, p. 117.

Duhem summarizes Rimini's view by saying that if God can, in a finite time divided into proportional parts, create an infinite magnitude by adding equal parts to each part of that time, he can also in same time divide a continuum into infinite proportional parts. Trying to accurately express the argument, he thinks that Rimini says that magnitude is a set of its parts, meaning that if the magnitude is given then all its parts are simultaneously given, too.

Duhem was not concerned with the way Rimini tried to solve the problem of actuality of the division of a continuum. Relying again on God's omnipotence, Rimini says that whatever part of the continuum you take, "and however small it is, it will [already] include parts which are actually divided in the divine conception. And just as in fact any continuum has infinitely many potential parts any one of which, however small, includes infinitely many [potential parts], and no part can be taken without it being divisible and without infinitely many such potential parts being taken, so I say that in God's conception the continuum is completely and actually divided into parts, of which any one is again completely actually divided and includes infinitely many actually divided [parts]".²⁹ It seems that, according to Rimini, the continuum can be divided into halves, then into thirds, and into quarters, and so on. If we remove God from the phrase "God's conception", we are still left with a conception. In this sense, we are closer to a Cantor-style theory of completed infinity, because the infinitely divisible continuum has already been 'traversed' by the definition of the principle of division.

Conclusion

Duhem's books on the history of science are extremely valuable, especially *Le Système du monde*. Even Alexandre Koyré, who does not accept Duhem's thesis of the organic connection between medieval thought and modern Galilean science, says that this book "is a work of permanent value whose richness of documentation is the fruit of labor so large that it confounds the mind; ... it remains a source of knowledge and an instrument of research which has not been replaced, and is therefore indispensable."³⁰ These words

²⁹ Gregory of Rimini, *op. cit.*, III, p. 224.

³⁰ Alexandre Koyré, "Review of Duhem's *Le Système du monde*", *Archives internationale d'Histoire des Sciences*, 35, 1956, p. 252.

still sound true, and Duhem's comprehensive reference to medieval texts is particularly useful because, even today, they are very difficult to access.

A pioneer of research in a particular field of science is not expected to produce a comprehensive text, especially not a multi-volume work. When it comes to our topic, Duhem's consideration of the issue of infinity in the Middle Ages, it must be noted that his work is not only the first of its kind, but is probably also the most valuable one.

Poincaré's epistemology of mathematics*

Poincaré had a peculiar position when it comes to understanding of mathematics. Without a doubt, one can trace Kant's line of thought in his work. He contested the logicist programme: the idea of reducing mathematics to logic. Usually we speak of Poincaré's appeal to intuition when it comes to arithmetic and of his conventionalism in geometry. A link can certainly be made between Poincaré's and Brower's position, even though, as opposed to Brower, Poincaré was not against modern logic in general. His idea was rather to show that inferences in mathematics cannot be reduced to logical inferences because they are essentially different. Mathematics requires more than logic: it relies on a special creative human ability, namely, *the intuition*. My aim in this paper will be to describe Poincaré's epistemological position with regard to mathematics. I will try to clarify two main points in Poincaré's position: the difference between logic and mathematics and the concept of intuition. I will put to the fore his special understanding of mathematical proof as *grasping the architecture*. An interesting metaphor Poincaré used to describe it is a game of chess. I will compare that understanding of mathematical proof with recent developments in game theoretical semantics and their application in the foundations of mathematics. It is interesting to see how this new turn reflects on Poincaré's position.

Poincaré against logicism

Let me start with a reflection on the place of logic in mathematics. Most of the interpreters underline Poincaré's disagreement with Russell and his link with Brower and the intuitionists. Both claims have indeed some justifications in Poincaré's thought.

* This article is a result of my research done within the project *Logico – Epistemological Foundations of Science and Metaphysics* (179067) supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

However, I agree with the often-overlooked remark from Mc Larty¹ that Poincaré had much appreciation for Frege's and Russell's "new logic" and he accepted it in a whole. That is the main difference in comparison with Brouwer and the later intuitionists. Poincaré was not what McLarty calls "a restrictive intuitionist" who rejects some parts of classical logic or mathematics. However, he did contest the project of reducing arithmetic to logic - the programme of logicism. According to Poincaré, inferences in mathematics cannot be reduced to logical inferences because they are not of the same kind. Mathematical inferences are different than those in logic: they rely on a special creative human ability called intuition.

There are two different interpretations of the nature of Poincaré's disagreement with the logicians. Brouwer and Goldfarb² find that intuition in Poincaré's philosophy of mathematics is a psychological term and that the key reason for his rejection of logicism was the lack of the psychological ground in the logicians' view of mathematics. Note that Frege's and Russell's idea was exactly to liberate the foundations of mathematics from psychological concepts. On the other hand, Detlefsen and Heinzmann³ argue that grounds for disagreement with the logicians was actually epistemological rather than psychological. I think that the later opinion is better justified and I will follow that interpretation in this paper. According to Detlefsen⁴, the conflict is not about the psychological origin of mathematical reasoning but rather about the question of whether the rational reasoning in general is homogeneous or not. The logicist programme assumed that it is homogeneous, so the same model of inference applies to mathematics as it does to logic. Poincaré, on the other hand, argued for the heteronomy of principles of rational thought and for the specificity of mathematical knowledge. This specificity did not drive him so far to claim that the logical reasoning has nothing to do with mathematics, as Brouwer did. Logic still keeps its ground as the important method of verification. However, when it comes to mathematical proofs, their nature is essentially different from that of logical proofs.

The difference in the view of the nature of mathematical knowledge appeared already between Leibniz and Kant. While Leibniz held that mathematics is entirely analytical in its

¹ Colin McLarty, "Poincaré & Logic & Intuition", *Philosophia Mathematica*, vol. 5, n° 3, 1997, pp. 97- 115.

² Pierre Brouwer, "L'œuvre philosophique" in *Henri Poincaré. L'œuvre scientifique, L'œuvre philosophique*, Paris, Alcan, 1914, pp. 205–264; Warren Goldfarb, "Poincaré against the Logicians", in *History and Philosophy of Modern Mathematics*, ed. W. Aspray, P. Kitcher, Minneapolis, Minnesota Press, 1985, pp. 61–81.

³ Michael Detlefsen, "Poincaré Against the Logicians", *Synthese*, vol. 90, n° 3, 1992, pp. 349 – 378; Gerhard Heinzmann, *Entre intuition et analyse: Poincaré et le concept de prédictivité*, Paris, Blanchard, 1985.

⁴ Detlefsen, "Poincaré Against the Logicians", p. 358.

character and that mathematical inferences are reducible to the laws of logic, Kant claimed that mathematical knowledge is essentially synthetic. Frege, Russell and Whitehead followed the first line of thought⁵ but Poincaré agreed with Kant on this point.

Mathematical induction

Poincaré opened his famous book *Science and Hypothesis*⁶ with a quite interesting observation concerning the nature of mathematics:

“The very possibility of mathematical science seems an insoluble contradiction. If this science is only deductive in appearance, from whence is derived that perfect rigour which is challenged by none? If, on the contrary, all the propositions which it enunciates may be derived in order by the rules of formal logic, how is it that mathematics is not reduced to a gigantic tautology?”

In this quote Poincaré pointed at the double nature of mathematics: on one hand, we associate a perfect rigour with it, but on the other, mathematics cannot be reduced to tautology because mathematical theorems enlarge our knowledge in such a way that we learn in the process more than it was contained in the axioms. Poincaré concluded that not all of mathematical knowledge is *analytical*. Therefore, Poincaré contested logicism. One of the most important arguments against it is the argument concerning mathematical induction.

Mathematical induction (or the proof by recursion) is one of Peano's axioms of elementary arithmetic, different from the other axioms. It is an "axiom — schema" which contains an infinity of axioms within it. It can be formulated in the following manner: If one theorem is true of 0, and if it is true of n , it is true of $n+1$, we can conclude that it is true of every natural number. Following Poincaré's analysis⁷, we have an infinity of syllogisms closed in one formula. If a property is true of 1, it is true of 2. If it is true of 2, it is true of 3. Every step is justified in a deductive way. Nevertheless, to make an infinite number of steps and to prove that the property applies to every natural number, we need

⁵ Although Frege thought that only arithmetic can be reduced to logic while geometry rests upon the intuition of space and thus is synthetic. Russell and Whitehead pushed the idea of logicism further and claimed that geometry is also reducible to logic.

⁶ Henri Poincaré, *Science and Hypothesis*, London, The Walter Scott Publishing Co., 1905.

⁷ *Ibid*, pp. 12-13.

more than syllogistic reasoning — we need an intuition. This typical mathematical reasoning thus shows that mathematics is not entirely analytical knowledge but also synthetic — it enlarges our knowledge behind what was contained in the premises. Mathematical induction enables us to bridge that crucial gap between the finite and the infinite, to climb from the individual to the general, without loss of necessity. However, this *gap* transcends logic! Every hypothetical syllogism in that axiom — schema is analytical and it can be verified, but the mathematical induction as a whole is synthetic.

The crucial distinction Poincaré made is the one between *verification* and *proof*. Verification is a process of analysis and it belongs to logic. In order to prove something in mathematics, however, we need the intuition. In the following section, I will elaborate more on the concept of intuition in order to explain how mathematical proofs differ from logical ones. Let me say in advance that definitions of mathematical entities, according to Poincaré, are not the expositions of their essential properties but rather the creations or constructions of those entities. There we can see again a close link with Kant. We need logic for verification and for precise exposition of what the intuition gave us in mathematics. It is also a mean for correction of intuition and the way to state its content more precisely because intuition, being a creative human capacity, can work somewhat more freely. However, the verification alone in mathematics is not sufficient. Poincaré states:

“Verification differs from proof precisely because it is analytical, and because it leads to nothing. It leads to nothing because the conclusion is nothing but the premises translated into another language. A real proof, on the other hand, is fruitful, because the conclusion is in a sense more general than the premises. The equality $2 + 2 = 4$ can be verified because it is particular. Each individual enunciation in mathematics may be always verified in the same way. But if mathematics could be reduced to a series of such verifications it would not be a science. A chess-player, for instance, does not create a science by winning a piece. There is no science but the science of the general. It may even be said that the object of the exact sciences is to dispense with these direct verifications.”⁸

The comparison of mathematical proof with a winning of a chess game is the important one, and I will come back to it later. Before that, let me touch upon the concept of mathematical intuition.

⁸ *Ibid*, p. 4.

Intuition

There are several studies of the concept of intuition in Poincaré's philosophy.⁹ It is usually said, with a good reason too, that Poincaré did not give one unique and clear account of that ability of the human mind and that he spoke of several different kinds of intuition.

Poincaré says:

“We have then many kinds of intuition; first, the appeal to the senses and the imagination ; next, generalisation by induction, copied, so to speak, from the procedures of the experimental sciences ; finally, we have the intuition of pure number [...], which is able to create the real mathematical reasoning. I have shown above by examples that the first two cannot give us certainty; but who will seriously doubt the third, who will doubt arithmetic?”¹⁰

If the first two, the intuition as appeal to the senses and the imagination¹¹, as well as the generalisation by induction are left aside, what is the real mathematical intuition ? I will start by listing different concepts of intuition described in the literature about Poincaré's philosophy of mathematics.

1. The conceptual intuition concerns the real, abstract mathematical objects. Mathematical objects are not available to the senses; mathematics is not about what is drawn on the board or about the written signs, but about ideal mathematical objects and their relations.¹²

2. The constructive intuition is the ability of arithmetical iteration — the ability of the human mind to understand that if one action is recognised as possible it can be repeated infinitely many times. This ability makes arithmetic possible because it is the root of

⁹ See for example Boutroux, "L'œuvre philosophique" ; McLarty, "Poincaré & Logic & Intuition"; Goldfarb, "Poincaré against the Logicians", Heinzmann, *Entre intuition et analyse: Poincaré et le concept de prédictivité* ; Janet Folina, "Poincaré on Mathematics, Intuition and the Foundations of Science", *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1994, pp. 217 – 226; Olga Pombo, "Conceptions of intuition in Poincaré's philosophy of mathematics", *Philosophy Study* (ISSN 2159-5313), vol.2, n°6, July-Aug 2012, pp. 384-397, etc.

¹⁰ Henri Poincaré, *The Value of Science*, New York, The Science Press, 1907. p. 20.

¹¹ Pombo argues in "Conceptions of intuition in Poincaré's philosophy of mathematics" that the sensible and the imaginative intuition play the role in the learning process of mathematics, both in ontogenetic and phylogenetic sense. Children need to see material things at first to begin to understand and the same was the case in the beginning of development of mathematics as a science.

¹² See *ibid.*

mathematical induction. Constructive intuition enables us to make the crucial step from finite to infinite.¹³

3. The architectonic intuition is the ability to grasp immediately the array of links in a deductive chain. It is the ability to grasp the architecture of a science that consists of a net of relations between propositions. This ability of *grasping the architecture of science* is a key to understand Poincaré's take on mathematical proofs, as I will elaborate a little later. As Rahman claims¹⁴, to construct a proof, we need to establish a connection between the structure or the architectonic of premises and the architectonic of a conclusion.

4. The foundational intuition is a foundational device for mathematics. It is the intuition that provides a full explanation of mathematical axioms. This intuition plays an important role in the choice of axioms in mathematics. A particularly interesting case is the choice of axioms of geometry. After the appearance of non-Euclidean geometries, Poincaré found it natural to describe axioms of geometry as conventions. However, those conventions are never completely arbitrary and we should investigate their origins. Poincaré states the following:

“The true geometrician makes this selection judiciously, because he is guided by a sure instinct, or by some vague consciousness of I know not what profounder and more hidden geometry, which alone gives a value to the constructed edifice. To seek the origin of this instinct, and to study the laws of this profound geometry which can be felt but not expressed, would be a noble task for the philosophers who will not allow that logic is all.”¹⁵

Olga Pombo¹⁶ compares this process with *abduction* as explained by Peirce. It is the art of choosing the best option between possible combinations and it demands certain creativity. Thus, a creative mathematician is guided by intuition in making choice of the axioms of geometry, having experience on one side and practical reasons on the other.

¹³ See Folina, "Poincaré on Mathematics, Intuition and the Foundations of Science" .

¹⁴ Shahid Rahman, "Remarks on Poincaré's Notion of Mathematical Rigour", *Logos & Episteme*, vol.1 (2010): 125-135.

¹⁵ Henri Poincaré, *Science and Method* (London: T. Nelson, 1914): 148-149.

¹⁶ Pombo, "Conceptions of intuition in Poincaré's philosophy of mathematics".

The architecture of science

After I have listed those different functions of the mathematical intuition, I would like to pay more attention to what I called the *architectonic intuition*. Detlefsen gave a brilliant analysis of Poincaré's epistemological position¹⁷ which offers an answer to the question of the specificity of mathematical proofs and of the reasons why they cannot be reduced to logical inferences. He provides an account of the meaning of *grasping the architecture* in mathematics. As I noted before, Detlefsen argues that the essential and we may say, the Kantian theme in Poincaré's rejection of logicism is the heterogeneity of rational thought; or, in other words, he claims that there is no unique model of inference. On that view, genuine mathematical inferences and logical inferences are different due to the fact that the way premises and conclusion are united is different in mathematics and in logic. Namely, an inference has three components: some premises, a conclusion and a *universal*, which expresses the architectonic of the theory to which the inference belongs by offering the model of comparison between the premises and the conclusion. According to Detlefsen, those universals are different in mathematics and in logic, which makes it impossible to reduce mathematical thinking to logical inferences:

“The difference between logical and mathematical inference, on this view, is one that centres on the choice of universals or architectures under which the premises and the conclusion are to be united: a mathematical inference being one that unites premises and conclusion as a ‘development’ under a mathematical architecture or theme, and a logical inference being one which unites premises and conclusion as a ‘development’ under logical architecture (if such there be).”¹⁸

This analysis apparently describes Poincaré's disagreement with the logicians as based on the epistemological reasons rather than psychological. The intuition in his philosophy of mathematics is not to be interpreted as a psychological concept but rather as a way to understand the connection between the structure of premises and the structure of conclusion in a mathematical proof.

This interpretation also clarifies previously mentioned tension described in the opening of *Science and Hypothesis* between the perfect rigour we associate with mathematics and its synthetic character. According to Detlefsen's analysis, the apparent

¹⁷ Detlefsen, "Poincaré Against the Logicians".

¹⁸ *Ibid.*, p. 362.

contradiction disappears once we understand that mathematical rigour does not imply that mathematical inferences have a logical character. Rather, in mathematics a different type of rigour is manifested. The fact that mathematics is not reducible to the laws of logic does not mean that its inferences are less certain. However, this mathematical rigour is a specific one and it relies on a specific architecture of the mathematical science. Mathematical intuition, which is crucial for mathematical proofs, also enables us to understand this architecture. As previously stated, to construct a proof, we need to establish a connection between the structure or the architectonic of premises and the architectonic of a conclusion.

Mathematics as a game

In this section, I want to develop the previously mentioned comparison Poincaré made between the mathematical proof and the game of chess. Poincaré writes the following:

“If you are present at a game of chess, it will not suffice, for the understanding of the game, to know the rules for moving the pieces. That will only enable you to recognize that each move has been made conformably to these rules, and this knowledge will truly have very little value. Yet this is what the reader of a book on mathematics would do if he were a logician only. To understand the game is wholly another matter; it is to know why the player moves this piece rather than that other which he could have moved without breaking the rules of the game. It is to perceive the inward reason which makes of this series of successive moves a sort of organized whole. This faculty is still more necessary for the player himself, that is, for the inventor.”¹⁹

This understanding of mathematics justifies indeed a label *intuitionist* or *pre-intuitionist* which is sometimes associated with Poincaré because it shows that his epistemological approach to mathematical proof is different to the purely formal one.²⁰

However, there was an interesting historical turn in the development of logic. By 1960, *dialogical logic* was developed by Paul Lorenzen and Kuno Lorenz, as a solution to some of the problems that arose in Lorenzen’s *Operative Logik* from 1955.²¹ Herewith, the

¹⁹ Henri Poincaré, *The Value of Science*, p. 22.

²⁰ See Shahid Rahman, Nicolas Clerbout and Radmila Jovanović, "The Dialogical Take on Martin-Löf's Proof of the Axiom of Choice", *South American Journal of Logic*, vol. 1, 2015, forthcoming.

²¹ The main original papers are collected in Paul Lorenzen and Kuno Lorenz, *Dialogische Logik*, Darmstadt: Wissenschaftliche Buchgesellschaft, 1978. Other papers have been collected more recently in Kuno Lorenz, *Logic, Language and Method: On Polarities in Human Experience*, Berlin-New York, De Gruyter, 2010, and

epistemic turn, initiated by proof theoretic approaches, was linked with the notion of games that provided the dynamic features of the traditional dialectical reasoning. Inspired by Wittgenstein's *meaning as use*, the basic idea of the dialogical approach to logic is that the meaning of the logical constants is given by the norms or rules for their use. Originally, dialogues were developed in the context of constructive mathematics and logic but since then they have served for comparison and combination of different logical systems.

Jaakko Hintikka combined the model-theoretical and the game-based tradition to create the Game Theoretical Semantics (GTS)²² that, like in the dialogical framework, grounds the concepts of truth or validity on the game-theoretical concepts, such as the existence of a winning strategy for a player, though, differently to the dialogical framework, it is build up on the notion of model. Hintikka claims that his semantic games are exact codifications of language — games in Wittgenstein's sense, at least if one accepts that the activities associated with quantifiers are *looking for* and *finding*. The inspiration came partially from Peirce who had already suggested in his second Cambridge Conference lecture held in 1898²³ that the difference between *every* and *some* can be explained in terms of who chooses an object. In the late fifties Henkin proposed that we could understand in game terms some sentences which did not have a satisfactory treatment in Tarski's semantics, such as those containing an infinite string of quantifiers²⁴. Henkin came up with the idea of a game for such a string of quantifiers between two players who are choosing the objects to replace the bound variables and he suggested that the truth conditions of such a sentence could then be expressed by its Skolem version. Hintikka generalized this idea to come up with a game theoretical semantics for classical, first-order logic, which is proven equivalent to Tarski's semantics, given that the axiom of choice holds. Hintikka

Kuno Lorenz, *Philosophische Variationen: Gesammelte Aufsätze unter Einschluss gemeinsam mit Jürgen Mittelstraß geschriebener Arbeiten zu Platon und Leibniz*, Berlin-New York, De Gruyter, 2010.

²² Jaako Hintikka, *Knowledge and Belief*, Ithaca, N.Y., Cornell University Press, 1962; Jaako Hintikka, *Logic, Language-Games and Information: Kantian Themes in the Philosophy of Logic*, Oxford, Clarendon Press, 1973; Jaako Hintikka, *The Principles of Mathematics Revisited*, Cambridge University Press, 1996; Jaako Hintikka, *Inquiry as Inquiry: A Logic of Scientific Discovery*, Dordrecht, Springer, 1999; Jaako Hintikka, Ilpo Halonen and Arto Mutanen, "Interrogative Logic as a General Theory of Reasoning", in Jaako Hintikka, *Inquiry as Inquiry: A Logic of Scientific Discovery*, Dordrecht, Springer, 1999, pp. 47–90. Rahman and Tulenheimo studied the relation between dialogical logic and GTS in Shahid Rahman and Tero Tulenheimo, "From Games to Dialogues and Back: Towards a General Frame for Validity", in: O. Majer and al. eds., *Games: Unifying Logic, Language and Philosophy*, Dordrecht, Springer, 2006, pp. 153–208.

²³ Charles Sanders Peirce, *Reasoning and the Logic of Things: The Cambridge Conferences Lectures of 1898*, ed. Kenneth Laine Ketner, Cambridge Mass., Harvard University Press, 1992.

²⁴ See Leon Henkin, "Some remarks on infinitely long formulas", *Infinitistic Methods, Proceedings of the Symposium on Foundations of Mathematics, Warsaw, 2-9 September, 1959*, 1961, pp. 167–183. See also Wilfrid Hodges, "Logic and Games" *The Stanford Encyclopedia of Philosophy* (spring 2013 Edition, URL = <http://plato.stanford.edu/archives/spr2013/entries/logic-games/>).

and Sandu came up with games with *imperfect information*²⁵, where the restriction can be made on the information available to the player so that he has to play sometimes not knowing (or forgetting) the previous moves made by the opponent. Those games yielded *Independence friendly first-order logic (IF logic)*, which has a greater expressive power than the classical, first-order logic since it can express more dependency patterns between quantifiers and other logical constants. It is expressive enough to enable formulating linearly and at first-order level, sentences containing branching quantification.

Interestingly, because of this great expressive power of IF logic, Hintikka proposed a new version of logicism, a conceptual reduction of mathematics to IF logic. His work has been contested on many levels, one of the reasons being that IF logic does not allow a complete proof system and it is thus inevitably related to model theory, which is in conflict with the idea of logicism. His realist position was also confronted with serious challenges.²⁶

However, the dialogical game theoretical tradition,²⁷ with its constructivist origin, comes closer to Poincaré's idea of describing a proof as a search for a winning strategy in a game. On one hand, it embodies the constructivist idea that to say that there is a proof one must be able to construct it; or in game-theoretic terms, the strategies must be *humanly playable*, which approximates Poincaré's insight. On the other hand, dialogical approach to meaning is very subtle because there are two kinds of rules for a dialogical game: *particle rules*, which are the rules of logical constants, and *structural rules*, which determine the general course of a dialogical game. The two kinds of rules allow us to distinguish between the *local* and *global* meaning, and thus to make the distinction between the level of play and the level of strategies, which captures Poincaré's point from the previously given quote

²⁵ Jaako Hintikka and Gabriel Sandu, "Informational Independence as a Semantical Phenomenon", ed. J. E. Fenstad, I. T. Frolov, R. Hilpinen, *Logic, Methodology and Philosophy of Science*, Amsterdam, Elsevier, 1989, pp. 571–589.

²⁶ For the study and the number of critiques of Hintikka's foundations of mathematics see Radmila Jovanović "Hintikka's Take on the Axiom of Choice and the Constructivist Challenge". *Revista de Humanidades de Valparaíso*, 2, 2013, pp. 135 – 152; and Radmila Jovanović, *Hintikka's Defence of Realism and the Constructivist Challenge* (PHD-Thesis, Lille/Belgrade): Université de Lille, forthcoming, 2015).

²⁷ A detailed account of recent developments since Shahid Rahman, *Über Dialoge, Protologische Kategorien und andere Seltenheiten*, Frankfurt-Paris-New York, Peter Lang, 1993, can be found in Shahid Rahman and Laurent Keiff, "On how to be a dialogician", ed. D. Vanderveken, *Logic, Thought, and Action*, Dordrecht, Kluwer, 2005, pp. 359–408; Laurent Keiff, "Dialogical Logic", *Stanford Encyclopedia of Philosophy* (Summer 2011 Edition, <http://plato.stanford.edu/entries/logic-dialogical/>) and Shahid Rahman, "Negation in the Logic of first degree entailment and *tonk*: A dialogical study", ed. S. Rahman, G. Primiero, M. Marion, *The RealismAntirealism Debate in the Age of Alternative Logics*, Dordrecht, Springer, 2012, pp. 213–250. For the underlying metalogic see Nicolas Clerbout, "First-order dialogical games and tableaux", *Journal of Philosophical Logic*, 2013 (doi: 10.1007/s10992-013-9289-z. URL <http://dx.doi.org/10.1007/s10992-013-9289-z>), pp. 1–17, and Nicolas Clerbout, *Étude sur quelques sémantiques dialogiques: Concepts fondamentaux et éléments de métathéorie* (PhD thesis, Lille 3 / Leiden: Universities of Lille and Leiden, 2013).

about the chess game. To understand the game of chess it is not enough to know how each of the pieces moves, one has to understand why a particular strategy should be chosen in order to win the game.

Described historical development of logic shows how interesting and important were Poincaré's insights concerning the nature of mathematical proof. He described it as *grasping the architecture* of science or grasping the structure of premises and a conclusion in a deductive chain through the constructive intuition and compared it with the game of chess. The constructivist and the game theoretical traditions embraced those ideas. On the other hand, the development of game theoretical approaches shows that logical, as well as mathematical inferences do not have to be sterile, as described by Poincaré, but they can display dynamic and epistemological features. A proof can be constructive and it can be built up on the notion of winning strategy in a game.

References

- Boutroux, Pierre (1914) "L'œuvre philosophique", in Henri Poincaré, *L'œuvre scientifique, L'œuvre philosophique*, Paris, Alcan, pp. 205-264.
- Clerbout, Nicolas (2013) "First-order dialogical games and tableaux", *Journal of Philosophical Logic*, doi: 10.1007/s10992-013-9289-z. URL <http://dx.doi.org/10.1007/s10992-013-9289-z>, pp. 1-17.
- Clerbout, Nicolas (2013), *Étude sur quelques sémantiques dialogiques: Concepts fondamentaux et éléments de metatheorie*, PhD thesis, Lille 3 / Leiden, Universities of Lille and Leiden.
- Detlefsen, Michael (1992) "Poincaré Against the Logicians", *Synthese*, 90 (3), pp. 349-378.
- Folina, Janet, (1994) "Poincaré on Mathematics, Intuition and the Foundations of Science", *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, pp. 217-226.
- Goldfarb, Warren (1985) "Poincaré against the Logicians", in *History and Philosophy of Modern Mathematics*, ed. W. Aspray, P. Kitcher, Minneapolis, Minnesota Press, pp. 61-81.
- Heinzmann, Gerhard (1985) *Entre intuition et analyse: Poincaré et le concept de prédictivité*, Paris, Blanchard.
- Heinzmann, Gerhard and David, Stump (2014), "Henri Poincaré", *The Stanford Encyclopedia of Philosophy* Spring Edition, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2014/entries/poincare/>.

- Henkin, Leon (1961) "Some remarks on infinitely long formulas", *Infinitistic Methods, Proceedings of the Symposium on Foundations of Mathematics, Warsaw, 2-9 September, 1959*, pp. 167–183.
- Hintikka, Jaako (1962) *Knowledge and Belief*, Ithaca, N.Y., Cornell University Press.
- Hintikka, Jaako (1973) *Logic, Language-Games and Information: Kantian Themes in the Philosophy of Logic*, Oxford, Clarendon Press.
- Hintikka, Jaako and Sandu, Gabriel (1989) "Informational Independence as a Semantical Phenomenon", ed. J. E. Fenstad, I. T. Frolov, R. Hilpinen, *Logic, Methodology and Philosophy of Science*, Amsterdam, Elsevier, pp. 571–589.
- Hintikka, Jaako (1996) *The Principles of Mathematics Revisited*, Cambridge, Cambridge University Press.
- Hintikka, Jaako (1999) *Inquiry as Inquiry: A Logic of Scientific Discovery*, Dordrecht, Springer.
- Hintikka, Jaako, Halonen, Ilpo and Mutanen, Arto (1999) "Interrogative Logic as a General Theory of Reasoning", in Jaako Hintikka, *Inquiry as Inquiry: A Logic of Scientific Discovery*, Dordrecht, Springer, pp. 47–90.
- Hodges, Wilfrid (2013) "Logic and Games" *The Stanford Encyclopedia of Philosophy*. Spring Edition, URL = <http://plato.stanford.edu/archives/spr2013/entries/logic-games/>.
- Jovanović, Radmila (2013) "Hintikka's Take on the Axiom of Choice and the Constructivist Challenge", *Revista de Humanidades de Valparaíso*, vol. 2, pp. 135–152.
- Jovanović, Radmila, *Hintikka's Defence of Realism and the Constructivist Challenge*, PHD-Thesis, Lille/ Belgrade, Université de Lille, forthcoming, 2015.
- Keiff, Laurent, "Dialogical Logic", *Stanford Encyclopedia of Philosophy*, Summer 2011 Edition, <http://plato.stanford.edu/entries/logic-dialogical/>.
- Ladyman, James (2014) "Structural Realism", *The Stanford Encyclopedia of Philosophy*, Spring Edition, URL = <http://plato.stanford.edu/archives/spr2014/entries/structural-realism/>.
- Lorenzen, Paul and Lorenz, Kuno, *Dialogische Logik*, Darmstadt, Wissenschaftliche Buchgesellschaft, 1978.
- Lorenz, Kuno (2010) *Logic, Language and Method: On Polarities in Human Experience*, Berlin- New York, De Gruyter.
- Lorenz, Kuno (2010) *Philosophische Variationen: Gesammelte Aufsätze unter Einschluss gemeinsam mit Jürgen Mittelstraß geschriebener Arbeiten zu Platon und Leibni*, Berlin-New York, De Gruyter.
- McLarty, Colin (1997) "Poincaré & Logic & Intuition", *Philosophia Mathematica*, vol. 5, n° 3, pp. 97–115.
- Peirce, Charles Sanders (1992) *Reasoning and the Logic of Things: The Cambridge Conferences Lectures of 1898*, ed. Kenneth Laine Ketner, Cambridge Mass: Harvard University Press.
- Poincaré, Henri (1905) *Science and Hypothesis*, London: The Walter Scott Publishing Co.,.
- Poincaré, Henri (1905) "Les Mathématiques et la Logique". *Revue de Métaphysique et de Morale*, vol.13, pp. 815–835.
- Poincaré, Henri (1907) *The Value of Science*, New York, The Science Press.
- Poincaré, Henri (1914) *Science and Method*, London, T. Nelson.
- Pombo, Olga (2012) "Conceptions of intuition in Poincaré's philosophy of mathematics", *Philosophy Study* (ISSN 2159-5313), vol.2, n° 6 (July-Aug), pp. 384–397.
- Rahman, Shahid (1993) *Über Dialoge, Protologische Kategorien und andere Seltenheiten*, Frankfurt, Paris – New York, Peter Lang.

- Rahman, Shahid (2010) "Remarks on Poincaré's Notion of Mathematical Rigour", *Logos & Episteme*, vol.1, pp. 125-135.
- Rahman, Shahid (2012) "Negation in the Logic of first degree entailment and *tonk*: A dialogical study", in: S. Rahman, G. Primiero, M. Marion eds., *The Realism-Antirealism Debate in the Age of Alternative Logics*, Dordrecht, Springer, pp. 213–250.
- Rahman, Shahid and Keiff, Laurent (2005) "On how to be a dialogician", in: D. Vanderveken ed., *Logic, Thought, and Action*, Dordrecht, Kluwer, pp. 359–408.
- Rahman, Shahid and Tulenheimo, Tero (2006) "From Games to Dialogues and Back: Towards a General Frame for Validity", in: ed. O. Majer and al. *Games: Unifying Logic, Language and Philosophy*, Dordrecht, Springer, pp. 153–208.
- Rahman, Shahid, Clerbout, Nicolas and Jovanović, Radmila (2015) "The Dialogical Take on Martin-Löf's Proof of the Axiom of Choice", *South American Journal of Logic*, vol. 1 (forthcoming).
- Worrall, John (1989) "Structural Realism: The Best of Both Worlds? ", *Dialectica*, vol. 43, pp. 99–124.

Philosophie de la physique aux XX^e et XXI^e siècles

Information et entropie : Ruyer, Merleau-Ponty, Atlan, Simondon

Le problème des relations entre information et entropie est bien oublié aujourd'hui, mais il a occupé dans la philosophie et dans les sciences en France, durant la seconde moitié du XX^e siècle – et pour être d'emblée plus précis : de 1950 à 1980 –, une position tellement centrale, qu'il pourrait suffire à désigner une époque (laquelle époque, dans la mesure où elle couple une réflexion sur le vivant à une interrogation sur la thermodynamique, donc sur l'éventuelle irréversibilité du temps, se laisserait volontiers nommer "post-bergsonienne"). L'origine historique de ce problème se trouve dans l'étonnement du mathématicien Norbert Wiener, lorsque celui-ci, dans son grand ouvrage *Cybernetics* (1948) – la cybernétique, qu'il a contribué à fonder, étant, rappelons-le, la science des machines comme telles, mais avec un accent mis sur les plus récentes d'entre elles, précisément les machines à information¹ –, constate que la formule de l' "information", définie comme contenu d'un message encodé par un émetteur et transmis à un récepteur qui le décode, est, en tant qu'elle se réduit au logarithme d'une probabilité, semblable à celle de l'entropie, à la seule différence du signe, la grandeur de l'information étant de valeur positive, celle de l'entropie étant de valeur négative. Ainsi, pour reprendre l'exemple classique, depuis Boltzmann, de l'entropie, lorsque les molécules d'un gaz parfait en viennent à être réparties homogènement dans une enceinte, cet état de la masse gazeuse étant le plus probable, il est celui qui nous apprend le moins sur le système formé par le gaz et l'enceinte, au sens où il est l'état le moins singulier, le moins remarquable, de ce système. Or, cette identité, au signe près, entre information et entropie, posée par Wiener, eut,

¹ C'est du moins la définition qui nous servira dans ce qui suit, puisque c'est celle que donne Ruyer (*La cybernétique et l'origine de l'information*, Paris, Flammarion [1954], coll. "Bibliothèque de philosophie scientifique", 2^e éd., 1967, p. 5). Mais Wiener est plus précis, en même temps que plus général, puisqu'il fait de la cybernétique la théorie de la *commande*, au sens des mécanismes destinés à en corriger d'autres ou à se corriger eux-mêmes (l'exemple paradigmatique, qui se ressent du contexte militaire où fut éprouvée la nécessité d'inventer une telle théorie, étant celui du missile à tête chercheuse). Le κυβερνήτης, c'est en effet, en grec, le pilote, celui qui commande un vaisseau, d'où "gouvernail", "gouverner", etc.

pendant la période considérée, la valeur essentiellement négative d'un interdit, et on peut considérer les quatre penseurs présentés ici comme ayant tenté, chacun à sa manière originale et irréductible, de franchir cet interdit. Il s'agit de Ruyer, de Merleau-Ponty, du biologiste et philosophe Henri Atlan, et de Simondon. Cette liste n'est pas exhaustive, mais des rapports internes, manifestés par des renvois explicites, se sont constitués entre les quatre doctrines, de sorte qu'elle possède une unité et une cohérence parfaitement attestées.

Raymond Ruyer parle très souvent de la cybernétique, mais il va jusqu'à lui consacrer un ouvrage entier en 1954 (remanié, de surcroît, en 1967), sous le titre *La cybernétique et l'origine de l'information*. C'est en effet en termes d' " origine " que se pose, pour Ruyer, la question de l'information, et il va jusqu'à donner, faisant ainsi violence aux règles oratoires de la *dispositio*, le titre " L'origine de l'information " au chapitre V de ce livre, chapitre qui acquiert ainsi le statut de partie totale, qui est effectivement décisif, et sur lequel devront donc se concentrer les analyses qui suivent. Car si, comme il est avéré, on doit distinguer entre l'usure thermodynamique à laquelle est soumise, en tant que processus utilisant de l'énergie, la transmission du message d'une part, et, d'autre part, le message lui-même, réputé inaltérable en tant que message, alors la cybernétique n'est-elle pas en train d'inventer, bien imprudemment, un " mouvement perpétuel de troisième espèce ", analogue à celui qu'ont rêvé, puis réfuté, la cinématique de Galilée (le frottement limitant la portée effective de l'inertie), puis la thermodynamique de Carnot (la dissipation de la chaleur rendant impossible sa reconversion indéfinie en travail) ? On voit que, dès à présent, la question de l'information accepte doublement une formulation thermodynamique. Mais surtout, le principe de la conservation de l'énergie (qui est, rappelons-le, le premier principe de la thermodynamique) impliquant que cette dernière, non seulement est incapable de se perdre, mais aussi, comme on le souligne moins souvent, de se créer – ce qui le place dans une relation de tension frontale avec le second principe de la thermodynamique, celui de l'entropie –, alors une grave difficulté se présente, pour l'énergie, quant à son " origine ", difficulté qui peut aussitôt être transposée au cas de l'information, et recevoir alors la formulation que voici : si l'information, du point de vue de sa quantité, ne peut, avec le temps, que décroître, alors comment concevoir, dans le passé, le surgissement d'un point optimum d'information ? Cette question, qui atteint finalement chez Ruyer une dimension théologique, peut se poser très simplement, de la manière suivante : puisqu'un émetteur radio ne peut que reproduire un message, et non pas l'inventer, alors comment concevoir qu'un autre émetteur radio se soit trouvé, éventuellement relayé par d'autres, à l'origine du

poème que l'on entend ? Ne faut-il pas, à cette origine, un être humain récitant le poème, peut-être même l'inventant ? C'est pour résoudre ce problème de l' " origine " de l'information que Ruyer forge, dans le chapitre V, l'appareil conceptuel qui constitue l'un des principaux apports philosophiques de son ouvrage.

Ruyer recourt à une distinction, elle-même très inspirée de Bergson et du troisième chapitre de *L'évolution créatrice*, entre deux sens du mot " ordre ", ce qu'il appelle l' " ordre homogène " d'une part (ainsi, " toutes les molécules d'eau du courant qui fait marcher le moulin vont à peu près dans le même sens "), et l' " ordre structurel " d'autre part (" un moulin et le corps du meunier ont une structure organisée ")² – là où Bergson parlait d' " ordre mécanique " et d' " ordre vital ". Or, affirme Ruyer, seul l'ordre dit " homogène " est incapable de s'accroître de lui-même, l'ordre " structurel " se caractérisant, précisément, par le fait qu'il peut s'accroître spontanément, sans violer le second principe de la thermodynamique. Une augmentation d'information, et partant son origination réitérée dans le réel, est donc compatible, du moins est-ce ce que veut Ruyer, avec ce dernier. Et dans ces conditions, l'information n'est plus le symétrique de l'entropie, puisqu'il existe une sorte d'entropie négative, par laquelle de l'ordre n'est pas seulement reproduit ou dégradé – comme c'est le cas de l'ordre " homogène " –, mais bien créé – l'ordre " structurant ". Et c'est l'apparition de structures répondant à un tel ordre " structurant " qui est accroissement d'information.

Mais alors, une conséquence radicale, et que d'aucuns pourraient estimer lourde à payer, s'ensuit de ce raisonnement : si, par définition, la formation de l'ordre structurant doit se penser à rebours de l'entropie, alors elle devra se concevoir hors de l'espace, et cette conséquence est parfaitement assumée par Ruyer, qui oppose lui-même au monde matériel un monde " trans-spatial ". Ce monde, que le reste de l'œuvre de Ruyer – ainsi *La genèse des formes vivantes* – baptise " domaine absolu de survol ", est le siège de forces immatérielles, telles que la conscience, et ce sont elles qui sont responsables de l'augmentation d'information.

Un premier modèle de solution au problème des rapports entre information et entropie, et une première manière de franchir l'interdit de Wiener – selon lequel on ne doit pas accorder à l'information le pouvoir de s'accroître –, apparaît dès lors avec Ruyer : puisque l'ajout d'information, en tant qu'il implique une inversion du courant de l'entropie, est impensable dans l'espace, alors c'est de l'espace qu'il faut sortir pour trouver l'origine

² Ruyer, *La cybernétique et l'origine de l'information*, p. 144.

de l'information (telle est, rappelons-le, la question spécifique dans laquelle Ruyer concentrait notre problème). On voit se dessiner en creux – disons-le pour poser une pierre d'attente – une solution différente du même problème, qui consisterait à dire que l'entropie est par elle-même, dans certains cas et par une vertu qui lui serait propre, créatrice d'information. C'est une solution que Ruyer admet parfaitement d'envisager, mais qu'il rejette non moins explicitement, en écrivant, dans un passage de *Néo-finalisme* qui définit l'entropie par la tendance des phénomènes à adopter le comportement le plus probable : “ Le hasard et la statistique [...] ne peuvent créer d'ordre. ”³

Merleau-Ponty, au moment où il décide de réorienter toute sa philosophie en direction d'une pensée de la nature, ne peut pas ne pas rencontrer, lui aussi, la cybernétique et le problème délicat de ses rapports avec la thermodynamique. C'est dans la partie datant de 1957-1958 du cours tenu au Collège de France sur la nature que ce problème est abordé. Le philosophe, résumant la position de Wiener et de la plupart des cybernéticiens, écrit : “ L'information, ce sera l'envers de l'entropie. L'information, c'est l'anti-hasard, la réalisation d'un état de haute structure qui n'est pas probable. L'information remonte une pente que l'entropie descend.⁴ ” Et pourtant, toujours si l'on suit Wiener, l'information, une fois surgie dans le monde par un processus qui, comme il vient d'être vu avec Ruyer, reste à expliquer, se trouve soumise au même principe conservatif que celui qui vaut pour l'énergie : ainsi, poursuit l'auteur, “ Une quantité d'information a été mise en circulation ; elle se dégrade ici ou là, mais dans l'ensemble elle se maintient, en tout cas elle ne s'invente pas ; tout au plus peut-elle se rétablir ”.⁵ Merleau-Ponty se trouve donc exactement en face du même problème que Ruyer, à savoir : comment, attendu que l'information une fois insérée dans le monde se dégrade et, partant, suppose à son origine un ou plusieurs points maximum d'information, se représenter le surgissement de celle-ci et sa cohabitation ultérieure avec la loi de l'entropie ? Et de fait, la seule référence livresque mentionnée par Merleau-Ponty dans les quelques pages consacrées dans ce cours à la théorie de l'information, c'est Ruyer et son livre sur l'origine de l'information. Or, ce que dit Merleau-Ponty au sujet du problème qui nous occupe, c'est, en substance, deux choses. D'une part, l'information se dégrade, sans qu'on puisse bien discerner, dans le texte qui nous a été laissé, si cette dégradation tient à l'information elle-même, ou seulement aux conditions de

³ Raymond Ruyer, *Néo-finalisme* (1952), Paris, PUF, coll. “ MétaphysiqueS », 2012, p. 296.

⁴ Merleau-Ponty, Maurice, *La nature. Notes du Collège de France (1956-1960)*, Paris, Seuil, coll. “ Traces écrites », 1995, p. 211.

⁵ *Ibid.*, p. 212.

sa transmission : “ L’information ne peut pas s’accroître, elle se dégrade : les bruits du téléphone font obstacle à la communication⁶. ” Mais, inversement – et c’est ici qu’on voit poindre la solution proprement merleau-pontienne au problème –, il y a une distinction fondamentale, du tout au tout, entre reproduire une information qui a déjà été inventée, et inventer l’information : “ Une machine peut indiquer automatiquement la loi d’une série de nombres en éliminant les anomalies. Il y a pourtant une différence avec la véritable invention, qui consiste à penser l’anomalie avec le reste⁷. ” Or, inventer l’information, cela ne peut être fait, selon Merleau-Ponty, que par un vivant : “ À partir de là, termine-t-il en effet, il est facile de voir comment la cybernétique tend à devenir une théorie du vivant et du langage⁸. ” L’argumentation merleau-pontienne reconduit donc à l’opposition de la machine et du vivant ; mais celle-ci, à son tour, renvoie, il faut le dire, à une conception philosophique bien plus profonde encore chez Merleau-Ponty – qui, ici encore, retrouve un grand geste ruyérien –, consistant dans la distinction, méthodologique mais aussi thétique, entre la pensée, philosophique et scientifique, qui reconstruit ou recompose, toujours après coup, un réel une fois donné (ce que *Le visible et l’invisible* appellera “ réflexion ”), et ce réel lui-même, saisi, si tant est qu’une telle saisie soit possible (le même ouvrage la nommera “ interrogation ”), à même le processus, unique et définitif, par lequel il en est venu, précisément, à se donner. C’est ainsi que l’on peut toujours recomposer, de l’extérieur et intellectuellement, le fonctionnement de la vision avec les parties de l’œil, mais cela ne nous re-donne en rien, si l’on peut dire, le processus concret par lequel la nature en est venue, au fil des générations et de ses tentatives aveugles, à faire un œil.

Merleau-Ponty, quatre ans après Ruyer et dans un sens bien proche de lui, énonce donc une nouvelle infraction à l’interdit de Wiener, en opposant, non plus cette fois un ordre “ structurant ”, créateur d’information, à un ordre “ homogène ”, où l’information ne peut que se dégrader, mais une manière *sui generis* et irrépérable, pour des êtres tels que le vivant, de venir à la présence, aux tentatives, toujours incomplètes, que la pensée, prolongée de la technique, peut effectuer pour répéter cette venue. Mais il reste, dans le cas de Merleau-Ponty (comme déjà dans celui de Ruyer), à statuer sur la nature de la force ou du processus par lequel l’information en vient à être inventée. La question tourne, de fait,

⁶ *Ibid.*

⁷ *Ibid.*

⁸ *Ibid.*

autour du caractère spatial ou extra-spatial de ce processus, et cette question, Merleau-Ponty refuse, du fait sans doute de sa massivité, de l'aborder de front dans le cours sur la nature.

Atlan, à partir d'une réflexion de savant sur l'organisation biologique, rompt complètement avec ces tentatives, qui, toutes plus ou moins, reviennent à faire de l'information une néguentropie, ce qui rend difficilement concevable sa possibilité physique. Son analyse, dans *Entre le cristal et la fumée* (1979), consiste tout entière, en effet, à distinguer entre deux sens de la redondance, comprise comme un caractère fondamental de l'entropie. Certes, la redondance est, le plus souvent, perte en information, puisque, précisément, au sein d'un phénomène qui se répète à l'identique, aucune des répétitions ne nous apprend davantage, sur ce phénomène lui-même, que les précédentes, si nombreuses soient-elles. C'est en ce sens-là que, selon Wiener, l'information est minimale, dans une enceinte renfermant un gaz, lorsque la répartition des molécules est parfaitement homogène, à savoir conforme à la plus haute probabilité. Mais la redondance peut, toutefois, représenter, comme telle, un surcroît d'information, par sa valeur de confirmation de l'occurrence d'un phénomène, et donc par la contribution qu'elle peut apporter à l'établissement d'une loi. On a ainsi affaire, chez Atlan, à la détection d'une particularité de l'entropie – en l'occurrence, la redondance – qui permettrait d'y loger la source de l'information, au lieu d'avoir à situer celle-ci dans des processus extra- ou néguentropiques, et donc virtuellement trans-spatiaux. Telle est l'idée principale d'Atlan en théorie de l'organisation, idée qu'il métaphorise, par deux fois, en imitant une formulation merleau-pontienne – “ l'organisation consiste précisément en une suite de désorganisations rattrapées ”⁹ – idée, par ailleurs, qui a souvent été baptisée “ information par le bruit ”¹⁰, mais qu'il convient, au-delà de ce slogan, de préciser.

Cette idée consiste, effectivement, à opposer, sur le fondement de la distinction des deux sens de la redondance, entre l' “ ambiguïté destructrice ” – défaut d'information tenant à la redondance du message : information-redondance – et l' “ ambiguïté-autonomie ” – défaut d'information tenant, cette fois, à son incomparabilité quasi totale : information-complexité¹¹ –, et, de là (presque comme Ruyer, et plus lointainement, presque comme Bergson), entre deux types d'ordre, l'ordre “ répétitif et redondant ”¹² d'une part, l'ordre

⁹ Atlan, Henri, *Entre le cristal et la fumée. Essai sur l'organisation du vivant* (1979), Paris, Seuil, coll. “ Points », série “ Sciences », 1986, p. 57 ; cf. p. 49, et Merleau-Ponty, Maurice, *La structure du comportement* (1942), Paris, PUF, coll. “ Quadrige », 1990, p. 6.

¹⁰ La formule est de l'important cybernéticien Heinz von Foerster.

¹¹ Atlan, Henri, *Entre le cristal et la fumée*, p. 64.

¹² *Ibid.*, p. 79.

vital, celui du hasard et du nouveau, d'autre part, Atlan préférant l'appeler "complexité". Mais alors, la question se pose de savoir si Atlan, en distinguant à son tour entre deux types d'ordre, ne retrouve pas à son tour la difficulté qui menaçait les entreprises de Ruyer et de Merleau-Ponty, à savoir, n'installe-t-il pas, implicitement, le "hasard" et le "nouveau", caractérisés comme facteurs d'organisation, au-delà de la sphère du spatialement déterminable ? Certes, Atlan évite toujours de poser frontalement la question de savoir si "hasard" et "nouveau" ont, sous sa plume, la signification ontologique qu'ils pourraient avoir chez Bergson, ou s'ils sont pourvus d'un simple sens thermodynamique¹³. Mais on peut dire que l'effort tout entier d'Atlan consiste à rendre informulable une interprétation spiritualiste de ses avancées, et cet effort, il le mène au moyen d'une affirmation selon laquelle le passage de l'ambiguïté destructrice à l'ambiguïté-autonomie, qui conditionne celui de l'ordre répétitif à la complexité véritable, tient, non pas à un fossé entre deux régions de l'être, mais à un changement de point de vue de l'observateur (de la partie au tout, de la voie de communication avec son *output* au système où elle se trouve elle-même située), lequel, en même temps, demeure certes enté sur le terrain de la physique, puisque, rappelle Atlan, il existe une "prédiction de Brillouin suivant laquelle, à partir du moment où des quantités énormes d'information seraient transmises, le facteur 10^{-16} ne serait plus suffisant pour les rendre négligeables en unités d'énergie".¹⁴ Ce qui signifie que le changement de point de vue, en tant qu'échange d'information, compte, lui-même, dans le bilan thermodynamique d'une réaction. Atlan tient énormément à cette hypothèse, sans laquelle son analyse tout entière est vulnérable, et dont le pendant physico-mathématique réside dans une "thermodynamique en réseaux", à savoir une thermodynamique déterministe et non probabilitaire, mais tenant compte, comme l'exige précisément toute thermodynamique, de la spécificité des "réseaux", avec leurs boucles de rétroaction.

Les recherches d'Atlan, vingt-cinq ans après le livre de Ruyer, représentent donc une percée considérable en direction d'une élucidation strictement physique de la réalité de l'information, puisque c'est à présent au sein de l'entropie elle-même, comme une de ses propriétés, qu'est mise au jour l'une des conditions de l'apparition d'information, cette condition étant la "redondance". Mais il reste à déterminer le statut exact – tient-il à notre représentation insuffisante, ou à l'essence intime de la réalité ?, vieille alternative

¹³ *Ibid.*, p. 81, note ; p. 82, note.

¹⁴ *Ibid.*, p. 125.

philosophique – du “ hasard ” et du “ nouveau ” qu’invoque Atlan, pour l’autre source de l’information.

La position de Simondon, exposée dans une thèse soutenue dès 1958, dédiée à la mémoire de Merleau-Ponty et dont le titre même renvoie à la notion d’information (*L’individuation à la lumière des notions de forme et d’information*), anticipe, par bien des côtés, celle d’Atlan, en particulier par le refus de s’adresser à une sphère “ trans-spatiale ” pour situer le point de jaillissement de l’information. Mais on peut même signaler que l’exemple favori des deux auteurs, précisément pour penser l’information, leur est commun, puisqu’il s’agit du cristal.

Une section tout entière du premier tome de la thèse de Simondon, paru en 1964 sous le titre *L’individu et sa genèse physico-biologique*, est consacrée à la résolution du problème wienerien des rapports entre information et entropie : c’est la section 4, intitulée “ De l’information à la signification ”, du deuxième point (“ Information et ontogenèse ”) du chapitre II (“ Individuation et information ”) de la deuxième partie de cet ouvrage, consacrée à “ L’individuation des êtres vivants ”. La récurrence du terme “ information ” dans les intertitres de Simondon est à elle seule révélatrice.

La solution de Simondon au problème de Wiener, et la manière simondonienne de répondre à l’interdit formulé par lui, sont, sur leur principe, identiques à celles d’Atlan : Simondon remarquait déjà que le surcroît d’information, dans un message donné, peut être apporté à la fois par l’imprévisibilité ou par l’unicité du phénomène encodé, mais aussi par sa prévisibilité ou sa répétitivité. Ainsi que l’écrit Simondon,

“Le signal d’information est en ce [premier] sens ce qui n’est pas prévisible, ce qui découpe le prévisible au point que l’énergie qui véhicule ce signal, ou les supports qui l’enregistrent, doivent avoir des états qui, à l’ordre de grandeur des signaux d’information (durée ou étendue selon le cas), peuvent être considérés comme prévisibles, pour que l’imprévisibilité des états du support ou de l’énergie modulée n’interfère pas avec celle du signal d’information”.¹⁵

Mais dans l’autre sens, “ l’information est ce qui implique régularité et retour périodique, prévisibilité. Le signal est d’autant plus facile à transmettre qu’il est plus facilement prévisible. ”¹⁶ Simondon distingue donc lui aussi, avant Atlan, entre une information-redondance et une information-imprévisibilité, ce qui se ramène à opposer, sur

¹⁵ Simondon, Gilbert, *L’individu et sa genèse physico-biologique*, Paris, PUF, coll. “ Épiméthée ”, 1964, p. 252.

¹⁶ *Ibid.*, p. 253.

le terrain de l'énergie elle-même et non pas sur un terrain qui lui serait transcendant, entre une dégradation qui serait perte d'information et une autre qui serait proprement gain d'information. Et dès lors, le maximum d'information sera contenu entre deux seuils, celui où le phénomène advenant, sans terme de comparaison possible, se trouve de ce fait pour ainsi dire muet, et celui où, identique à tous les autres, il ne nous apprend rien non plus sur sa spécificité, mais pour les raisons inverses (Simondon donne, à cet égard, l'exemple de la plage de sable, où tous les grains sont identiques, du moins à l'échelle du regard humain non prolongé d'un instrument¹⁷). De ce point de vue donc, la solution simondonienne au problème posé par Wiener présente les mêmes qualités d'élégance, car d'économie, que celle d'Atlan, qui en sera si proche.

La distinction, tracée par Simondon, entre les cas où la répétition est informatrice, et ceux où, comme le voulait effectivement Wiener, elle produit simplement un brouillage de l'information, est plus claire encore dans le cours sur la communication professé en 1970-1971, repris dans le volume *Communication et information* (paru en 2010) :

“La marche lente d'un défilé, quand elle devient plus tendue, plus serrée, fait sentir que le danger d'affrontement se précise ; le bruit social n'est pas seulement un bruit de fond ; il amorce une information et suscite une motivation qui annoncent la communication éthologique. [...] Contrairement à ce qui se passe dans les communications échangées par l'intermédiaire d'un canal technique, le bruit de fond n'est donc pas toujours le contraire du signal significatif ; même aléatoire, le bruit de fond joue déjà un rôle de stimulation et de synchronisation”.¹⁸

Dans ce passage encore, on a affaire, presque, aux deux moments de l'idée qui sera développée ultérieurement par Atlan : il y a, premièrement, une information par le bruit, et cela même lorsque, deuxièmement, le bruit n'a pour règle que le hasard.

Les solutions de Simondon et d'Atlan diffèrent cependant sur un point crucial, et c'est la mise en valeur de celui-ci qui, dans la mesure où elle soulignera une originalité de Simondon peu aperçue jusqu'à présent et donc riche de potentialités pour la philosophie et la science aujourd'hui (ce qui justifie notre choix de placer Simondon au terme de cette revue), doit être notre tâche ici. Car si le geste théorique accompli par Atlan consiste à opérer sur le message lui-même, et à se demander s'il véhicule une information-redondance ou une information-complexité, celui de Simondon implique, tout au contraire, que l'on se place aux pôles de ce message, et réside, quant à lui, dans un basculement des perspectives,

¹⁷ *Ibid.*, p. 252-253.

¹⁸ Simondon, Gilbert, *Communication et information*, eds. Nathalie Simondon et Jean-Yves Château, Paris, Les Éditions de la Transparence, coll. “ Philosophie ”, 2010, p. 95.

depuis le pôle de l'émetteur jusqu'à celui du récepteur. L'opposition pertinente devient, à présent, celle du signal "émis" ou "transmis" au signal "reçu".¹⁹ C'est en effet que

"Les signaux sont d'autant mieux transmis qu'ils se confondent moins avec l'uniformisation du prévisible ; mais pour qu'ils soient reçus, pour qu'ils s'intègrent au fonctionnement d'un système, il faut qu'ils présentent une analogie aussi parfaite que possible avec ceux qui pourraient être émis par le dispositif récepteur si on l'utilisait comme émetteur ; il faut qu'ils soient presque prévisibles".²⁰

Or – et c'est ici que la plus grande originalité de Simondon est perceptible, en même temps que la radicalité des réformes de pensée auxquelles cette originalité nous contraint –, la capacité de l'information à être reçue est, selon l'auteur, foncièrement relationnelle, en aucun cas absolue. C'est de la même manière que le relief, dans la vision binoculaire (selon un exemple cher à Simondon, qui trouve un de ses emplois dans le contexte présent), consiste dans la relation entre les deux images, lesquelles, d'ailleurs, doivent être, point trop différentes, mais pas non plus tout à fait identiques.²¹ Voilà pourquoi, surtout, l'aptitude à être transmise ne constitue en rien une "qualité" de l'information, tout au plus une "eccéité"²² (selon le terme de Simondon) de celle-ci, mais en vérité une relation entre l'émetteur et le récepteur, relation dont la nature se laisse enserrer par le terme "signification" : "Les signaux doivent rencontrer pour être reçus des *formes préalables* par rapport auxquels ils sont *significatifs* ; la signification est relationnelle."²³ La découverte du pouvoir informateur de la répétition trouve donc, chez Simondon, sa condition ultime dans l'ontologie de la relation soutenue par ce philosophe, et dans la substitution qu'il entend opérer de celle-ci, à l'ancienne ontologie de la substance, ou de la matière in-formée.

Le problème posé par Wiener, et la stricte relation de symétrie établie par lui entre les deux termes de ce problème – information et entropie –, constitue donc la pierre de touche à laquelle on peut mesurer, dans la période considérée (les décennies 1950, 1960 et 1970), la proximité et la distance des quatre penseurs que nous avons pris en considération : Ruyer, Merleau-Ponty, Atlan et Simondon.

Pour briser cette relation de symétrie, en effet, une première option est possible, c'est celle qui consiste à poser, en sens inverse de l'entropie, une tendance à l'organisation,

¹⁹ Simondon, Gilbert, *L'individu et sa genèse physico-biologique*, p. 254-255.

²⁰ *Ibid.*, p. 254.

²¹ *Ibid.*, p. 255-257.

²² *Ibid.*, p. 255.

²³ *Ibid.*

au non-probable, qui serait la force productrice d'information elle-même. Cette force, Ruyer la conçoit comme “ ordre structurant ”, par opposition à l’ “ ordre homogène ”, et il la détermine en dernière analyse comme conscience. Merleau-Ponty, dans un sens tout différent, y voit le mouvement même de la venue à l'être, à l'écart de toutes les reconstructions que la pensée réfléchissante en donne, sans s'apercevoir (le plus souvent) qu'elle le suppose. Le danger propre de cette première option, mieux paré sans doute par Merleau-Ponty que par Ruyer, consiste à hypostasier la réalité de l'information dans un monde psychique se dérochant à la prise expérimentale, et par là même à la mesure.

La seconde option consiste à installer la possibilité de la production d'information dans l'entropie elle-même, et non plus à l'extérieur d'elle, dans un pouvoir paradoxal qui lui est propre, et qui tient à la redondance qu'elle porte en elle : il y a une information par la redondance, qui certes ne constitue pas toute l'information, mais qui donne néanmoins sa part à la répétition dans la formation d'organisations aussi compliquées que les êtres vivants. Mais alors, on est mis au défi de tracer, au sein du monde physique et non plus entre lui et son au-delà, une frontière entre les phénomènes soumis à la redondance entropique d'une part, et, d'autre part, l'efficace du “ hasard ” et du “ nouveau ” (pour reprendre le vocabulaire d'Atlan), ce qui oblige – et l'affaire n'est pas mince – à élucider le statut ontologique de ces deux derniers termes. D'où, chez Atlan, le recours à la prédiction de Brillouin et à une nouvelle théorie physique pour l'accompagner – la “ thermodynamique en réseaux ” –, et, chez Simondon, le passage à une “ ontologie de la relation ”, où la signification du message, laquelle est une pure relation entre l'émetteur et le récepteur, a cependant, et précisément en tant que relation, rang d'être.

Universalité des systèmes complexes

Dans le labyrinthe des systèmes complexes, où chaque cas représente un problème particulier posé par la Nature, il est difficile de trouver le fil d'Ariane qui les relierait et permettrait de trouver la sortie. D'un point de vue épistémologique, on peut dire que ce fil est découvert dans l'introduction d'un paramètre d'ordre, spécifique pour chaque système. Ceci a permis non seulement la comparaison des transitions ordre-désordre dans des cas différents, mais aussi l'application des lois d'échelle pour démontrer l'universalité des phénomènes physiques dans des systèmes variés.

Dans cet exposé, nous allons essayer de donner quelques idées sur cette approche universelle, due à un scientifique hors de l'ordinaire, Pierre-Gilles de Gennes, qui a su expliquer le comportement et l'analogie entre les systèmes complexes, comme supraconducteurs, cristaux liquides, ou polymères en interaction, par des considérations simples, faisant appel à la théorie des transitions de phase de Landau et au groupe de renormalisation.

L'oeuvre de P.-G. de Gennes a été récompensée par de nombreux prix et médailles, dont la couronne a été le prix Nobel en Physique pour 1991.

Universalité : approche épistémologique ou scientifique ?

L'universalité des systèmes complexes peut être discutée dans le cadre d'une approche épistémologique ou purement scientifique. Dans cet exposé, nous choisissons la première approche, sans oublier qu'il est impossible de la séparer complètement de l'approche scientifique.

Parmi les différents systèmes complexes, il y a des exemples dans le domaine de la matière condensée, comme les magnétiques et les supraconducteurs solides, et dans le domaine de la matière molle, comme les cristaux liquides, les polymères, les milieux poreux, les grains de sables, les bulles de savon, etc. Pour démontrer ce qu'il y a d'universel

dans des systèmes très différents, nous considérons supraconducteurs, cristaux liquides et polymères.

La sortie de labyrinthe du comportement complexe de ces systèmes est donc due à Pierre-Gilles de Gennes. Son approche est dans l'esprit du grand physicien russe Lev Landau, qui a été le premier à démontrer que les transitions de phase dans des systèmes très différents présentent des caractères universels. Dans beaucoup de cas, le fil d'Ariane qui a mené à l'universalité a été la théorie de Landau des transitions de phase. Ceci implique l'existence d'un paramètre d'ordre, caractéristique pour chaque cas particulier, et permet d'établir une analogie entre les cas différents. Mais la notion de paramètre d'ordre s'est avérée importante même en l'absence de transition de phase, comme chez les polymères en solution. Dans ce dernier cas, de Gennes a découvert qu'en introduisant un vecteur de paramètre d'ordre, avec $n=0$ composantes, on trouve une analogie complète entre les transitions de phase, par exemple en système magnétique dont le paramètre d'ordre est à n composantes, et les polymères en solution. A première vue, il semble que l'existence d'un paramètre d'ordre à zéro composantes est paradoxale. En fait, ceci traduit une caractéristique très importante des chaînes de polymères en solution – marche aléatoire auto-évitante. Ces découvertes ont eu des conséquences très importantes en physique statistique, et en physique en général. Afin d'illustrer l'universalité des systèmes complexes, nous allons nous limiter aux deux cas suivants :

- analogie entre les transitions de phase en supraconducteurs et en cristaux liquides ;
- application aux problèmes de conformation des polymères en solution.

Ce choix est en partie inspiré par le commentaire du comité Nobel, fait lors de l'attribution à Pierre-Gilles de Gennes du prix Nobel de Physique 1991 “ pour la découverte que les méthodes développées pour étudier les phénomènes d'ordre dans les systèmes simples peuvent être généralisées à des formes plus complexes de la matière, en particulier aux cristaux liquides et polymères ”, commentaire assez restrictif devant l'importance de la contribution de Pierre-Gilles de Gennes à toute la physique.

D'un point de vue épistémologique, pour comprendre comment le savoir est atteint, il est important de remarquer que l'universalité de phénomènes très différents ne peut être mise en évidence sans une connaissance parfaite des cas particuliers. C'est la condition nécessaire pour trouver les analogies. Aussi, il est nécessaire de maîtriser tout l'appareil de physique théorique, et en particulier celui de la statistique moderne : application des lois d'échelle, changement du nombre des dimensions, groupe de renormalisation, etc.

Analogie entre supraconducteurs et cristaux liquides

Avant d'aborder le sujet, il faut remarquer une chose qui est de première importance pour le progrès de la science, et par conséquent, pour l'épistémologie. C'est l'échange libre des informations scientifiques. Ceci est très bien illustré par l'histoire de la recherche en supraconductivité. Le travail expérimental et théorique s'est développé séparément à l'Ouest et à l'Est. Après la découverte, dans les années 1930, des alliages métalliques supraconducteurs par Lev Shubnikov, c'était en Union Soviétique le sujet principal de la recherche en ce domaine, menée par Piotr Kapitsa, Lev Landau, et Shubnikov lui-même. Ce dernier a été mis en prison durant les purges de Staline, et plus tard exécuté en 1945. Le brillant théoricien Lev Landau, qui a été aussi mis en prison pendant un an par Staline, a formulé une théorie générale de transitions de phase continues. Avec Vitaly Ginzburg, il l'a appliquée en 1950 à la supraconductivité. Cette théorie, tout en restant phénoménologique, a permis de comprendre plusieurs propriétés des supraconducteurs. A cause de la guerre froide, les physiciens aux Etats-Unis n'ont appris ces développements que beaucoup plus tard. Les journaux de physique soviétiques ne pouvaient pas traverser le "rideau de fer". Entre-temps, les physiciens américains Bardeen, Cooper et Schrieffer ont formulé la première théorie microscopique du phénomène. Ici, le rôle de De Gennes est tout à fait singulier. Ayant appris le russe pour pouvoir lire les articles scientifiques soviétiques, il a compris qu'il fallait combler l'écart entre la théorie microscopique BCS et la théorie phénoménologique Ginzburg-Landau, ce qu'il a fait en même temps que le théoricien russe Gor'kov à Moscou.

Avant de s'attaquer aux deux classes de matériaux très différents que sont les supraconducteurs et les cristaux liquides, de Gennes avait déjà l'expérience de la recherche dans le domaine du magnétisme et de la supraconductivité. Dans les années 50, il a travaillé sur les magnétiques pour sa thèse de doctorat en France et pendant son stage de post-doctorat à Berkeley, USA. Sa curiosité scientifique, qui ne l'a pas abandonné de toute sa vie, l'a ensuite amené à se tourner vers le sujet de la supraconductivité.

Au-dessous d'une température critique, T_c , les matériaux supraconducteurs ont des propriétés étranges, comme le passage d'un courant électrique sans résistance et l'expulsion du flux magnétique. Conséquence des corrélations fortes entre les électrons,

ces propriétés s'expliquent par l'hypothèse qu'il s'agit d'un phénomène quantique macroscopique.

Dans ce cas-là, une seule fonction d'onde Ψ décrit l'ensemble des paires d'électrons condensés : c'est le paramètre d'ordre dans la théorie phénoménologique de Ginzburg-Landau des transitions de phase. Les propriétés d'un supraconducteur sont déterminées par le rapport de deux longueurs caractéristiques. La longueur de cohérence ξ donne l'échelle de la variation du paramètre d'ordre, et l'épaisseur de pénétration λ donne l'échelle de la variation des quantités électromagnétiques.

Si le rapport $\kappa = \lambda/\xi > 1/\sqrt{2}$, en présence d'un champ magnétique H la transition de l'état supraconducteur vers l'état normal se passe au voisinage d'un champ critique supérieur. La transition de phase est du 2nd ordre, et le paramètre d'ordre tend vers zéro d'une manière continue. Pour les champs plus faibles qu'un champ critique inférieur, le flux magnétique est expulsé du volume supraconducteur, c'est l'état Meissner. Dans le domaine entre ces deux champs, la pénétration partielle du flux se fait sous la forme d'un réseau régulier de tourbillons de flux parallèles : c'est l'état mixte, appelé aussi "phase Abrikosov-Shubnikov", d'après le théoricien et l'expérimentateur qui l'ont découvert. Pour $\kappa = \lambda/\xi < 1/\sqrt{2}$, la transition du 1^{er} ordre se fait directement entre l'état supraconducteur et l'état normal.

Pour comprendre comment un comportement similaire dans les cristaux liquides a été expliqué par de Gennes en utilisant l'approche de Landau, il faut d'abord voir de quoi il s'agit. Expérimentalement, on trouve des propriétés intrigantes. Leurs molécules adoptent un degré d'ordre, par exemple ils pointent tous dans une même direction. Cet "ordre désordonné" peut être NEMATIQUE (quand les molécules sont plus ou moins parallèles, mais leurs centres de masse sont distribués au hasard, comme dans les liquides), SMECTIQUE (quand les molécules, bien qu'alignées, forment des couches superposées, mais à l'intérieur de chaque couche il n'y a pas d'ordre à grande distance), ou bien cholestériques (avec les molécules arrangées en hélice), que nous n'allons pas discuter ici, voir Fig. 1.

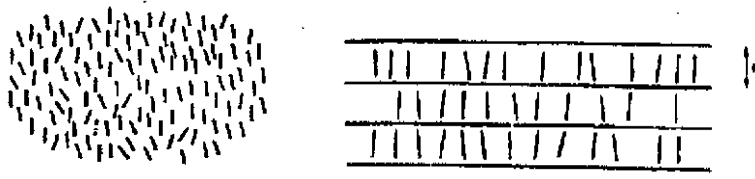


Fig. 1. A gauche : distribution des molécules dans un nématique, à droite : distribution des molécules dans un smectique.

Inspiré par un résultat expérimental, de Gennes a supposé que les smectiques (du type A) peuvent être décrits par une fonction d'onde Ψ similaire à celle des paires de Cooper dans les supraconducteurs. L'amplitude de ce paramètre d'ordre représente la modulation de la densité du cristal liquide, et sa phase donne la position des couches. Cette analogie a conduit à la prédiction qu'au voisinage d'une transition smectique A $\Rightarrow$ nématique (du 2nd ordre) la température de transition T_c baisse si une distorsion de torsion ou de flexion est imposée. Ces distorsions sont l'analogue d'un champ magnétique pour un supraconducteur. T_c diminue si h , le "champ magnétique" correspondant, est différent de zéro. Poussant l'analogie encore plus loin, de Gennes a démontré l'existence de deux types de transition dans les smectiques, juste comme pour les supraconducteurs, ainsi que l'existence d'une nouvelle phase, analogue à l'état mixte. Cette phase, trouvée expérimentalement beaucoup plus tard, contient un réseau régulier de dislocations, ce qui permet de relaxer la déformation. Pour cela, il a été nécessaire d'introduire un paramètre équivalent $\kappa = \lambda / \xi$. Dans le cas présent, la longueur de cohérence donne l'échelle de la variation de l'amplitude du paramètre d'ordre, imposée par une perturbation locale ; par exemple, ξ est le rayon du "cœur" des dislocations dans l'état mixte.

L'épaisseur de pénétration λ mesure la pénétration de la distorsion imposée à la surface du spécimen. Le résultat est une frappante ressemblance des diagrammes de phase des supraconducteurs et des cristaux liquides.

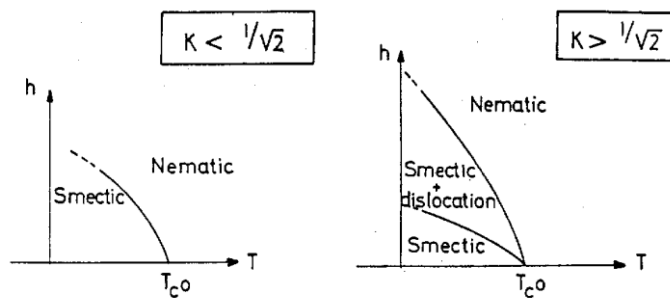


Fig. 2. Diagramme de phase de la transition smectique A-nématique. h mesure l'amplitude de la déformation appliquée. Pour la transition supraconducteur-métal normal, $h \rightarrow H$, le champ magnétique extérieur, et dans l'état mixte, on a la pénétration de tourbillons de flux.

Le problème du volume exclu : des gaz opalescents aux polymères

Au début des années 70, les deux problèmes principaux de la physique statistique étaient les fluctuations critiques de certains gaz denses, et la conformation des polymères en solution. En régime critique, ces gaz présentent de grandes régions de faible densité en alternance avec des régions de densité élevée, ce qui leur donne une apparence opalescente. Les polymères sont des macromolécules en forme de chaîne, dont les joints peuvent se tordre quand ils sont en solution dans un liquide approprié. Une telle chaîne présente un type de marche aléatoire, dont l'extension spatiale devrait varier comme la racine carrée de sa longueur N . Pourtant, expérimentalement, la loi de variation est plus rapide.

Le comportement de ces deux systèmes tellement différents s'explique par un seul phénomène : celui du volume exclu. Les molécules du gaz, de même que les sous-unités des chaînes des polymères, ne peuvent pas s'entrecouper. Le problème des lois des puissances particulières pour la croissance des fluctuations critiques des gaz a été résolu par la méthode puissante de groupe de renormalisation, inventée par K. Wilson en 1972. (Pour cette découverte, Wilson a obtenu le prix Nobel de 1982.)

D'un point de vue épistémologique, il est important de souligner que la nouvelle méthode peut s'appliquer à des cas plus généraux, comme les systèmes de même symétrie. Un exemple est donné par la transition d'un métal vers l'état ferromagnétique, où l'on

regarde les fluctuations du paramètre d'ordre qui est le vecteur de la magnétisation locale. Par analogie, on peut traiter les fluctuations d'un paramètre d'ordre général avec un nombre arbitraire de composantes n (pour les gaz, $n=1$, et pour les magnétiques, $n=3$).

La grande découverte de De Gennes était d'établir un parallèle entre ce dernier cas (paramètre d'ordre à n composantes) et les polymères en interaction. Dans les deux cas, le volume exclu se présente sous forme de marche aléatoire auto-évitante. Il faut donc calculer les diagrammes représentant les classes des configurations s'entrecroisant, pour lesquels il faut faire la correction à la contribution de toutes les configurations. De Gennes a remarqué que l'expression mathématique de ces diagrammes pour polymères était identique à celle correspondant à un paramètre d'ordre vecteur à n composantes, si on choisissait une valeur appropriée de n .

Quelle était cette valeur magique ? Il s'est avéré que c'était $n=0$! Le sens physique de cette découverte est clair : bien que la notion du paramètre d'ordre à zéro composantes paraisse paradoxale, elle décrit le problème du volume exclu. La découverte de $n=0$ a eu des conséquences importantes. Avant, la symétrie de renormalisation était considérée comme une caractéristique exclusive des transitions des phase avec brisure de symétrie accompagnante, ce qui n'est pas le cas chez les polymères. Par la suite, plusieurs autres cas de $n=0$ ont été découverts, avec des applications nombreuses en physique statistique.

Pour retourner aux propriétés globales des polymères, il suffit d'utiliser la correspondance avec les transitions de phase. Pourtant, il faut faire deux remarques.

La première se rapporte au nombre des dimensions de l'espace d . Ici, une observation de Wilson a été cruciale. Il a remarqué que pour $d > 4$ la divergence des diagrammes reliés aux configurations s'entrecroisant disparaît. Donc si on suppose que la dimension de l'espace est proche de 4, on peut calculer les exposants critiques en effectuant une expansion en petit paramètre $\varepsilon = 4 - d$. Pour passer aux trois dimensions, il suffit de prendre ensuite $\varepsilon - 1$.

La seconde remarque se rapporte à une autre contribution importante de De Gennes à la physique statistique des polymères. C'est l'introduction des lois d'échelle. Il s'agit de considérer le système à une échelle assez grande pour que toutes les propriétés deviennent auto-similaires.

Cela peut être illustré à l'exemple d'un polymère en solution semi-diluée. Même avant de Gennes, on savait que dans ce cas la chaîne se met sous forme de rouleau du rayon moyen $R \approx aN^\nu$. Ici, le coefficient a dépend de la nature du polymère et du solvant. Par

contre, à l'aide des transformations d'échelle, on peut démontrer que l'exposant ν est universel, il a la même valeur pour tous les rouleaux dans les solutions tridimensionnelles. L'idée de base est de prendre une chaîne qui a les monomères de dimension a et la constante d'interaction adimensionnelle u , qui comprend les interactions monomère-monomère, monomère-solvant, et solvant-solvant. Groupant g monomères en sous unités nouvelles, voir Fig. 3, on recalcule la dimension a_1 et la constante d'interaction u_1 . L'itération prochaine donne les sous-unités de g^2 monomères de dimension a_2 et de constante d'interaction u_2 , etc. Cette procédure est répétée jusqu'à ce que la constante d'interaction ne change plus. Dans ce régime auto-consistant, on peut démontrer que R suit une loi de puissance universelle $R \approx \text{const. } N^\nu$, CQFD. La valeur de l'exposant ν , et les autres propriétés des solutions de polymères, peuvent alors être calculées par la méthode du groupe de renormalisation.

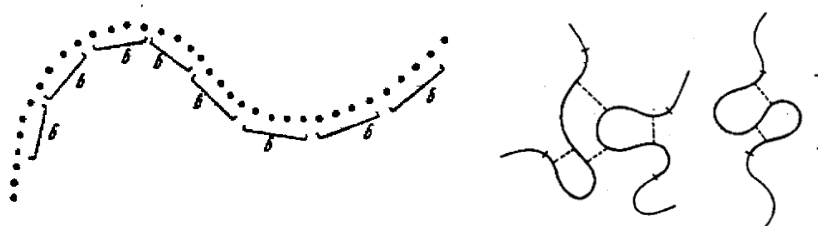


Fig. 3. *A gauche : groupement de g monomères successifs en nouvelles sous-unités. A droite : nouvelles sous-unités en interaction, de dimension indiquée par la parenthèse.*

En général, l'approche des transformations d'échelle consiste à établir une loi à petite échelle, à partir des paramètres qui contrôlent le système (la plus grande difficulté est de déterminer ces paramètres correctement), à reconstruire une échelle au-dessus, et à résoudre le problème par une analyse dimensionnelle.

Cette idée de “ *scaling* ” n'est pourtant pas nouvelle, elle était déjà connue dans l'Antiquité. Platon l'avait formulée dans la *République*, où il proposait d'examiner la nature de la justice et de l'injustice d'abord au niveau de l'Etat et ensuite au niveau de l'individu, partant du plus grand au plus petit, et en les comparant.

De Gennes lui-même avait comparé cette approche à celle des impressionnistes. Avec le progrès de la photographie, l'attention au détail n'était plus intéressante pour les peintres. Il y avait d'autres choses à explorer, ce qui a mené à des mouvements comme l'impressionnisme.

Conclusion

En conclusion, les deux exemples que nous avons discutés nous montrent – chacun à sa manière – l'universalité des systèmes complexes.

L'analogie entre cristaux liquides et supraconducteurs nous montre une étonnante mais puissante corrélation entre des objets qui n'ont rien de commun, mais obéissent à des lois identiques. Cette analogie remarquable, qu'on a pu trouver en introduisant un paramètre d'ordre approprié dans chaque cas, montre que des lois universelles s'appliquent dans des domaines tout à fait différents de la physique.

La reconnaissance du fait que les polymères sont une réalisation du groupe de renormalisation a permis de comprendre beaucoup mieux la nature de la symétrie de dilatation. À part les polymères, qui ont ouvert la boîte de Pandore, on trouve la réalisation des états avec zéro ($n=0$) composantes du paramètre d'ordre dans de nombreux exemples d'objets géométriques stochastiques. Les états quantiques désordonnés, les processus de croissance stochastiques, la dynamique chaotique, ils possèdent tous la symétrie de dilatation, universelle dans tous ces systèmes complexes.

D'un point de vue épistémologique, il faut noter que pour démontrer un comportement universel une approche globale est nécessaire. En remplaçant l'approche semi-empirique par une approche rigoureuse on peut trouver des lois simples, en accord avec l'expérience et indépendantes de la nature de système particulier. En même temps, cela a permis de comparer des systèmes complexes divers et de trouver des analogies possibles.

Scaling Down The Space*

All science is cosmology, I believe.

K. Popper

Prologue

Two principal world outlooks may be traced down to the pre-Socratic Greece: (i) the atomistic (Abderian) and the fractal (Clazomenian) schools. The principal proponents of the former were (possibly) Leucippus, Democritus and Epicures, whereas it was Anaxagoras who proposed a multileveled material world, introducing the concept of self-similarity.¹ The latter has now been recognized as a fractal structure. In fact, it was Plato who first, as far as we know, mentioned the concept of self-similarity, as described in his *Politeia*, 5009 D.² As noticed by Daniel Funaro,

“This seemingly matches the experienced space with its ruling imagination of a three-dimensional Euclidian Space. However, Plato’s space as universe holds more than the result of ordering axioms that served as a feasible arrangement for Newton’s inertial frame. In fact all spaces are contained in each other forming one just seemingly stable universe like a Hausdorff Space. Though, based on its self-motined origin, this agglomerated space is more comprehensive, it is not static, but expanding like our observed universe. Three-dimensional time d^3 measures the continuous expansion of three-dimensional space d^3 caused by the underlying self-motion like the frequency of a calibrated clockwork. “He (God) made an eternal image, moving according to number, even that which we have named Time” (Timaeus 37d).”³

* I am much indebted to my wife Ljiljana for her invaluable help and support. This paper is a result of my work supported by the Ministry of Science and Environmental Protection of Serbia, under the contract No ON171020.

¹ Petar V. Grujić, “The Concept of Fractal Cosmos: I. Anaxagoras’ Cosmology”, *Serbian Astronomical Journal*, Belgrade, n° 163, 2001, pp. 21-34.

² Werner Jaeger, *Die Formung des griechischen Menschen*, Berlin-New York, Walter de Gruyter, 1973.

³ D. Funaro, arXiv:0906.1874v1 [physics.gen-ph] 10 Jun 2009.

From then on two principal cosmological lines were characteristic for the European world outlook, as variations on Democritus' and Anaxagoras' themes. Democritus' concept of atoms and void and an infinite universe, both in space and time, was adopted by Epicures, Giordano Bruno, Isaac Newton and others, whereas Anaxagoras' idea of a hierarchical structure of the material world had to wait the 20th century to be applied in many areas, from microscopic level to cosmology.⁴ Immanuel Kant made use of both concepts when describing the structure and the dynamics of his cosmology. Though devised as *αρχέ* in the presocratic sense, Anaxagoras' and Democritus' fundamental assumptions were of different nature. For Democritus, the only physical reality consisted of atoms and void, the matter being made out of the former, in all possible varieties, whereas Anaxagoras layed down a fundamental principle, rather than entities.⁵ Today, we recognize his hierarchical structuring as an application of the principle of *symmetry*, more precisely, of one of possible symmetries, that of *scaling transformations*.⁶ It includes in part the Democritian concept of matter, with a particular structuring, but leaves intact the other Democritus' entity, the *void*. However, modern research of the structure of space-time has opened new possibilities concerning the structure of empty space. It has turned out that by applying the scale transformations on space-time itself, new fundamental properties of our world emerge.⁷

Anaxagoras' laid down the principle Παντι παντος μοιρα ενεστι (in everything there is a portion of everything), which today has been recognized as referring to *fractal structures*.⁸ It corresponds to a multilayered world, with a geometric form repeating at each level, with layers nested indefinitely one into the other, like Russian dolls. This Platonic approach puts stress on the geometric form of the material constituents of the world, with a pattern repeating at each subsequent level of smaller dimension. As exposed in *Timaeus*, Plato conceived the elementary constituents to be regular geometric figures, each endowed with one of possible symmetries. It is this symmetry that repeats at any level, and scales down from one level to the next one. The geometric form will play a significant role in Descartes' description of the nature of the material world, as its essential attribute.⁹

⁴ Petar V. Grujić, "The Concept of Fractal Cosmos: II. Modern Cosmology", *Serbian Astronomical Journal*, Belgrade, n° 165, 2002, pp. 45-65.

⁵ Ch. Mugler, "Le problème d'Anaxagore", *Revue des études grecques*, 69, 1956, p. 314.

⁶ Petar V. Grujić, "The Concept of Fractal Cosmos: III. The Present State", *Serbian Astronomical Journal*, Belgrade, n° 182, 2011, pp. 1-16.

⁷ Nottale Laurent, *The Theory of Scale Relativity, Fractal Space-Time, Nondifferentiable Geometry and Quantum Mechanics*, Imperial College Press, London, 2012.

⁸ B. Mandelbrot, *The Fractal Geometry of Nature*, New York, Freeman, 1983.

⁹ Alexandre Koyré, *Études Galiléennes*, Paris, Hermann, 1980; *Galileo Studies*, Brighton, Harvester 1977.

Symmetry, transformations and scaling

Symmetry plays significant roles in arts, music and aesthetics in general.¹⁰ Symmetry we observe in Nature, as well as those constructed in mathematics, in geometry in particular, play very important role in science. The most important conservation laws are derived from the symmetry one finds in mathematical expressions and it is the transformations we carry out on physical systems which reveal the physical quantities which are preserved. As worked out by Emmy Noether, the law of energy conservation stems from the invariance of time translation, conservation of the momentum from the space translation, and the angular momentum from the rotational invariance. Topology, as worked out by Henri Poincaré, shows how the abstract mathematical space is transformed under various mathematical transformations. One important attribute of abstract space is the so-called connectivity, which counts the ways subspaces can be brought to coincide. For instance, a sphere is a single connected surface, since any of its closed curves may be made circular by a continuous transformation, and any two circles may be fused by continuous transformations. In the case of a torus (donut-shaped), however, one may draw two circles, for instance, which can't be fused without leaving the surface, and it is, therefore double connected. Abstract spaces are thus classified by their topological properties and one of the principal questions we are facing in observational and theoretical cosmology is the problem of finding out what is the topology of our universe. Fractal systems possess their own so-called *fractal dimension* D_f , which is different from the topological dimension.

Scaling transformations, which preserve the shape, but diminish or extend the size of the object, are symmetry preserving and thus of particular importance in science, considering the importance of symmetry itself, as we mentioned above. In statistical physics (the so-called *renormalization group*), classical few-body problem,¹¹ etc. scaling transformations appear to be a powerful tool for evaluating many physical quantities.

¹¹ Petar V. Grujić, "Classical Calculations of the Near-Threshold Multiple-Escape Processes", *Comm. At. Mol. Phys.* 18, 1986, pp. 47-74.

Fractal structures

Mandelbrot gave the following definition of fractal structure: Objects, curves, functions or sets are fractal when “their form is extremely irregular and/or fragmented at all scales”.³ If the same pattern repeats at every scale, one talks about self-similar fractals. Fractal structures, in particular the self-similar ones, exact or approximate, abound in nature, science and arts (see, e.g. Mandelbrot 1983). Here we mention cosmology, molecular biology (DNA), linguistic (Chinese letters), paintings (Pollock), music (Josquin des Prez), optics, field theory (Casimir effect), superconductivity, solid state physics, and science of material, to mention some of them only. In Nature one finds fractals as snowflakes, plants, herbs, sea coasts, bronchial system in our lungs, turbulent flow of liquids, bacterial colonies, etc. Some of them appear to be exact fractals, like snowflakes, the other only approximate. Fractals with infinite number of levels exist only in mathematics, otherwise this number is finite.

It is interesting to note here that Anaxagoras’ *σπερματα* (seeds) find their counterpart in the *Upanishads*:

[6.12.2] ”Verily, my dear, that finest essence which you do not perceive - verily, my dear, from that finest essence this great Nyagrodha [sacred fig] tree thus arises. [6.12.3] ”Believe me, my dear, that which is the finest essence - this whole world has that as its soul. That is Reality [satya]. That is Atman. That art thou, Svetaketu.”¹²

Here the role of *vous* is played by the (individual) soul (Atman), as an integral part of the Brahman (World’s soul). This ”finest essence” we recognize today as genetic material, which determines our phenotype. Hence, to *Upanishads*’ *Tat twam asi* (That are thou) corresponds a particular fractal level.

¹² Robert Ernest Hume (Ed). Chandogya Upanishad, in *The Thirteen Principal Upanishads*, Oxford, Oxford University Press, 1975.

Scaling the space-time

Considering our primeval dichotomy of Abderian *versus* Clazomenean paradigm, it is the atomic (discrete) structuring of the micro world which prevents scaling towards the infinitely small. Atoms are divisible, but subatomic particles, like electrons, mesons, quarks, etc are not. As for the higher (cosmic) levels, it is the empirical (observational) evidence which will decide the number of higher levels.

In the case of self-similarity, three scale domains can be distinguished (i) the quantum (microworld, where scale invariance does not apply), (ii) the classical (mesoscopic) domain (invariant) and (iii) the megascopic (cosmic) domain (invariant?).

The lower level of indivisibility of matter is determined by the laws of Quantum mechanics, which is itself based on the existence of lower limits of the corresponding physical quantities. The most fundamental of them is Planck's constant h , which appears as an absolute unit of the angular momentum. We recall that the latter is conserved in an isolated physical system and corresponds to the rotational motion which in its turn appears absolute in space, as the famous Newton's experiment with bucket and water demonstrates. The existence of h implies the minimum physical length, the so-called Planck length $\ell_P = (\sim G/c^3)^{1/2}$. Democritus' *indivisibility*, on the other hand, does not refer to ordinary matter, but to the measure of a physical quantity *angular momentum*, and by implication to the length. For the portions of space which have dimensions smaller than ℓ_P , the standard space attributes, like metrics, cease to exist and new ones emerge. We imagine the space attributes at this level to fluctuate in an aleatory manner, which has inspired some theoreticians to call it *space foam*. Fractality as such loses sense, at least the standard one. In its turn, fractal objects have attributes which appear strange to our normal experience. So, fractal curves are nondifferentiable, at least within the Newton-Leibniz integro-differential calculus. The length of the fractal curve between two points separated by a finite length in the Euclidean sense, turns out to be infinite, for instance.

Much information can be obtained by making symmetry transformations of space-time, the exchange of positions of the bodies within physical systems, into an exchange of physical attributes of elementary particles (such as charge and „colours“) *etc.* Symmetry turns out to be a powerful guiding force, and symmetry transformations a fertile resource of new physical properties, laws and rules. What can one gain from symmetry transformations of the scale of space-time, scaling transformations? Rather surprising, it

turned out that these transformations have a great explanatory and innovatory power. In the recent comprehensive monography by Laurent Nottale,³ from Meudon, who is the leading authority on the subject, has presented the latest results he and other researchers have produced making use of the scaling transformations.

We shall list some of the most interesting results here. Those who are acquainted with physics at least at the informative level, know what somebody has put as: God gave us two equations only – Newton's and Schrödinger's. These are the cornerstones of the classical and Quantum mechanics respectively. Millions of people are engaged all over the world in solving some variants of these equations. Since they are fundamental laws of Nature, they cannot be derived from other expressions. In particular, Schrödinger's equation cannot be derived from classical theory. However, it turns out that one can derive it from the classical expression by applying scaling transformations (space dilation) on isotropic space,¹³ provided that Heisenberg uncertainty conditions are satisfied. One thus establishes *quantum scale-laws*, in analogy with the standard quantum laws.

Applications are carried out in Quantum mechanics, classical and quantum electrodynamics, vast areas of physics, both quantum and Newtonian, cosmology, biology, celestial mechanics (planetary system), to cite just some of the very diverse applications. Scaling transformations turned out to reveal another dimension of the abstract physical space, which is directed to the „depth“ of space-time and provide a new insight into physical reality.

Reality and illusion

Acquiring knowledge about the external world consists of an interplay of empirical evidence and our mental constructs (theoretical models). We are aware of our „ability“ to make mistakes in modeling reality, but the empirical evidence is usually taken as reliable experience. In order to acquire sense data, however, we have to make use of our senses. The most frequently used means is visual inquiry. This, however, can be misleading, as the so-called optical illusions demonstrate. They may be common ones and thus provide a general distorted image of the external world. The case in point is the geocentric

¹³ L. Brenig, „Is quantum mechanics based on an invariance principle?“, *Journal of Physics A: Mathematical and Theoretical*, vol. 40, n° 17, 2007, pp. 4567-4584.

cosmology, which kept the mankind with the wrong outlook for millennia. It is not an accident that such collective illusion appears in astronomy and cosmology in a broader sense, since the observations are the only means of getting the information from the outer space. It is, therefore, of the utmost importance to have a correct interpretation of the observational evidence, which amounts to constructing an adequate cosmological model. What is the situation in this context today?

The current version of the so-called standard model of cosmology assumes an isotropic, globally homogeneous, expanding universe. However, there are models based on other paradigms, like the so-called fractal model, which argues for the hierarchical cosmology in an anaxagorian sense. It makes use of the current evidence for the clustering of galaxies, which first form local clusters, then clusters, superclusters (clusters of clusters) etc. It is essentially a non-uniform distribution arrangement. Two remarkable features of this model should be stressed here: (i) the average cosmic density of matter is zero, (ii) the distribution appears isotropic. The latter means that the universe looks the same in every direction to any observer who watches the outer space from a constituent unit. We conclude that the uniform distribution is not a necessary condition for observational isotropy. Since the cosmic isotropy is firmly established for the time being, it is the estimate of the space distribution of the material content which is to decide whether the universe is better described by the Standard model, rather than by the hierarchical (fractal) one. But it is not the only dilemma cosmologists are facing at present.

A cosmologist from Meudon, Jean-Pierre Luminet, has developed a peculiar model based on an assumption about the cosmic topology.¹⁴ He assumed that the cosmic space can be described by a multiply-connected manifold, the so-called *wraparound* universe. In such a model, light from a distant galaxy reaches us along many different paths, so that what we observe may be a multiple optical version of the same cosmic object. It is as we were in a room with mirror walls, floor and ceiling, with light rays reflecting many times before entering our eyes. As a consequence, we live in a universe which is much smaller than our observational evidence shows. It is not an optical illusion in a proper sense, but a direct consequence of a particular topology of the cosmos we live in.

Both cases we mentioned, that of the hierarchical and of the wraparound cosmological models, demonstrate two things relevant to the epistemological aspect of the

¹⁴Jean-Pierre Luminet, *L'univers chiffonné*, Paris, Fayard, 2001. English translation by E. Novak, *The Wraparound Universe*, Wellesley, MA, AK Peters, 2008.

cosmological issue:¹⁵ (i) empirical evidence is important, but not a reliable guide to the truth, and (ii) when we acquire data only via a single sense, theoretical models play the crucial role in deciphering the information we are making use of.

Another point regarding the reliability of the empirical evidence must be made too. Our observations (both visual and otherwise) are accurate at the mesoscopic level within the technical limits of our observational tools. Examining the structure of a physical object we observe irregularities, depending on the “resolution power” of the tool, including naked eye. What appears a smooth surface, for instance, may turn out to be rough, after “zooming” the instrument. Generally, we assume to be able to reach the resolution high enough to be satisfying “for all practical purposes”, unless we happen to deal with a fractal object. Once we conclude this to be the case, further attempts to achieve an “absolute precision” are doomed to failure. Fortunately, as we have seen above, all fractals in Nature have a final number of nested levels.

¹⁵ Petar V. Grujić, “Some Epistemic Questions of Cosmology”, *Foundations of Science*, 12, no 1, 2007, pp. 39-83.

Philosophie de la biologie aux XX^e et XXI^e siècles

Ciprian Jeler
Département de Recherches interdisciplinaires
dans le Domaine sciences Sciences Socio-Humaines
Université Alexandru Ioan Cusa
Iași, Roumanie

Normativité biologique et ‘ niveaux ’ de la vie.

Remarques sur le rapport de Canguilhem à la théorie de l'évolution

1. Introduction

Dans un article consacré au “ style français ” en épistémologie, Jean-François Braunstein¹ affirme que l’une des caractéristiques principales de l’épistémologie française est celle de refuser de distinguer, en ce qui concerne les sciences, entre ce qu’une tradition anglo-saxonne appelle le “ contexte de découverte ” et le “ contexte de justification ”. Une telle distinction, affirme ce commentateur, permet aux épistémologues anglo-saxons de laisser de côté – ou bien de lui accorder un rôle secondaire ou accessoire – le contexte de découverte et de se concentrer – comme à leur véritable objet – au seul contexte de justification, à savoir aux théories et surtout à la hiérarchie des arguments et des données expérimentales ou empiriques qui soutiennent ou qui sous-entendent une certaine manière de traiter un champ de phénomènes donné. De son côté, le refus des épistémologues de style français d’accorder un rôle décisif à cette distinction rendrait compte de l’importance de l’historicité pour l’épistémologie française, à savoir du fait que, pour une certaine tradition française, l’épistémologie est constitutivement inséparable d’une histoire des sciences. D’une manière plus directe, cela nous introduirait dans un type de rapport des philosophes à la science pour lequel, selon l’expression de Bachelard, faire de l’épistémologie revient à s’occuper de l’histoire des sciences, mais en même temps, cette histoire des sciences devient du même coup une historiographie de la vérité.²

D’une manière analogue, mais, cette fois-ci concernant plus directement George Canguilhem, Pierre Macherey distingue entre l’histoire des savants ou l’histoire telle quelle

¹ “ Bachelard, Canguilhem, Foucault. Le ‘ style français en épistémologie ’ ”, dans *Les philosophes et la sciences*, dir. Pierre Wagner, Paris, Gallimard, 2002, pp. 920-963.

² Gaston Bachelard, *Le matérialisme rationnel*, Paris, PUF, 1953, p. 86.

est vue par les scientifiques et la “ vraie histoire ” des sciences, telle qu’elle est, pour Macherey, pratiquée par Canguilhem :

“Alors que le chemin de l’histoire réelle *va du concept au phénomène à travers deux médiations étroitement solidaires : expérimentation et théorie*, l’histoire vue par les savants *se fonde sur une conception hiérarchique des niveaux, de l’observation à la théorie*, qui permet à la fois des opérations de substitution (phénomène = concept = théorie) et la conception de l’histoire comme un *enchaînement* des théories : on part des théories, – et on en reste aux théories –, qu’on relie les unes aux autres parce qu’elles constituent *l’élément le plus achevé* de la pratique scientifique. Démarche typiquement idéaliste.”³

Sans entrer dans les détails de cette discussion, il est important de noter que cette distinction entre l’histoire des savants et “ la vraie histoire ” se formule en fonction de la question des niveaux. Les scientifiques qui font l’histoire de leur discipline – ou plutôt l’histoire de tel problème de leur discipline – parlent normalement de la hauteur de l’état présent de ce problème, parlent, autrement dit, rétrospectivement, à partir de la manière dont un problème est résolu aujourd’hui pour remonter ensuite vers des préfigurations ou approximations antérieures de cette résolution dans le passé. Mais ce n’est pas tout, car c’est seulement parce qu’ils instituent dès le départ une distinction nette entre des observations et des théories, une distinction de niveaux donc, qu’ils arrivent à jeter ce type de regard rétrospectif sur leur discipline ou leur domaine, et c’est la position initiale de cette distinction de niveaux qui leur permet d’envisager des thèses antérieures comme étant des préfigurations de la résolution actuelle. C’est à partir de cette distinction de niveaux que les scientifiques arrivent à éliminer l’historicité de l’histoire de leur discipline, car, aux dires de Macherey, dans le cadre d’une telle histoire, je cite, “ aucune durée n’[...] est plus possible ”⁴ ou bien “ au début et à la fin on doit retrouver la même chose, *entre les deux il ne se passe rien* ”⁵. C’est en refusant de poser cette distinction initiale entre les observations et les théories qu’une histoire des sciences telle qu’elle est pratiquée par Canguilhem peut se concentrer sur les concepts, et peut suivre le rapport du concept au phénomène étudié à travers deux médiations, pourtant interdépendantes, qui sont l’expérimentation et la théorie. C’est donc en refusant de *partir* d’une distinction de niveaux que, aux dires de Macherey,

³ Pierre Macherey, “ La philosophie de la science de Georges Canguilhem ”, dans *La Pensée*, nr. 113, février 1964, p. 61.

⁴ *Ibid.*, p. 59.

⁵ *Ibid.*

Canguilhem substitue la filiation des concepts à l'enchaînement des théories.⁶ C'est autrement dit, le refus d'une séparation nette entre les niveaux de la théorisation scientifique qui introduit l'historicité proprement dite dans l'histoire des sciences.

Mais ces points indiquant le rapport entre l'historicité et la question des niveaux ne visent que la pratique théorique de Canguilhem et non le contenu de ses thèses. Il s'agit plutôt de considérations de méthodologie, d'une certaine manière de faire l'histoire des sciences, et non d'une évaluation des conclusions théoriques et surtout philosophiques que Canguilhem arrive à tirer de ses études sur différents champs ou notions que l'histoire des sciences nous offre. Notre propos ici visera justement à essayer d'envisager un aspect du rapport entre l' " historique " et la question des niveaux dans les thèses mêmes que Canguilhem avance, et cela non pas sur la manière dans laquelle il faut ou il faudrait faire l'histoire des sciences, mais sur une question plus précise et moins méthodologique, à savoir sur la thèse canguilhemienne de la normativité biologique. Ajoutons tout de suite que cette question de la normativité biologique se conjugue, chez Canguilhem, selon deux lignes directrices et non pas une seule. D'un côté, la normativité biologique est appelée par Canguilhem pour décrire le rapport entre la santé et la maladie, et cela arrive dans les trois textes importants que Canguilhem consacre à ce problème, à savoir l'*Essai sur quelques problèmes concernant le normal et le pathologique* de 1943, l'article " Le normal et le pathologique " publié en 1951 et repris en 1952 dans *La connaissance de la vie*⁷, et, enfin, les " Nouvelles réflexions concernant le normal et le pathologique " qui datent de 1966, et qui ont été intégrés à l'édition de *Le normal et le pathologique*⁸ que nous lisons et utilisons aujourd'hui. Pourtant la normativité biologique ne s'applique pas seulement au rapport de la santé avec la maladie, mais elle reçoit chez Canguilhem un sens beaucoup plus vaste pour autant que cette notion est appelée aussi à donner lieu à une considération sur la vie " en général ", et par cette " vie en général " nous avons en vue l'aspect ou la théorie la plus englobante que la biologie moderne nous présente, à savoir la théorie de l'évolution. Chacun des trois textes qui s'attache au problème du normal et du pathologique passe, comme par une sorte de passage obligatoire, par des considérations visant la théorie de l'évolution et la manière dans laquelle la notion de normativité biologique pourrait y trouver sa place. C'est comme si la normativité vitale dans le rapport santé-maladie et la normativité vitale concernant la théorie de l'évolution étaient des questions jumelles, mais c'est le sens

⁶ *Ibid.*, p. 61.

⁷ Paris, Hachette, 1952 (abrégé ici CV).

⁸ Paris, PUF, " Quadrige ", 1966 (abrégé ici NP).

de cette parenté qui nous reste caché et qui, à notre avis, ne pourrait pas être résolu sans une étude compréhensive portant sur l'œuvre entière de Canguilhem.

Sur ce point, rien ne serait, à notre avis, plus dangereux que d'essayer de répondre d'une manière hâtive à la question " pourquoi ", à la question de savoir ce qui explique cette co-apparition constante de la question du normal et du pathologique, d'une part, et de la question d'une qualification ou d'une interprétation de la théorie de l'évolution, de l'autre part. Pourrions-nous dire que Canguilhem essaie de fonder ses thèses sur le normal et le pathologique en les appuyant sur cette théorie englobante qu'est la théorie de l'évolution d'inspiration darwinienne ? Ou bien, pourrait-on dire au contraire qu'il ne s'agit, dans cette contiguïté essentielle des deux problèmes, que d'une tentative de Canguilhem d'étendre ses thèses sur la santé et la maladie à un autre domaine biologique – la théorie de l'évolution – mais que cette extension serait à penser comme analogique plutôt que comme exprimant des tendances fondatrices ? Pourrait-on enfin soutenir que Canguilhem voudrait fonder à la fois le rapport entre le normal et le pathologique et la théorie de l'évolution sur une notion métaphysique de normativité vitale qui ne surgirait pas d'une réflexion appliquée et précise sur des données scientifiques ? Voici trois manières d'expliquer cette contiguïté essentielle des deux problèmes, trois manières qui ont certes l'avantage d'être facilement concevables mais qui, à notre avis, ne sont ni soutenues par les textes de Canguilhem, ni défendables – mais, au contraire, critiquables – à partir de ces textes mêmes.

Nous nous garderons donc ici, pour notre part, de céder à cette tentation philosophique de répondre à la question " pourquoi ". Notre but dans ce texte sera beaucoup plus modeste et il pourrait être exprimé d'une manière très concise : il s'agira de montrer que, même si les positions de Canguilhem sur la normativité biologique qui est impliquée dans le rapport entre la santé et la maladie ne semblent pas changer d'une manière significative entre les trois textes dans l'espace d'une vingtaine d'années, il y a pourtant des glissements conceptuels qui affectent l'autre versant de cette notion, à savoir le rapport de la normativité biologique à la théorie de l'évolution. Il y a, à notre avis, des différences dans la manière canguilhemienne d'envisager la théorie de l'évolution dans ces trois textes, et nous allons essayer de les exposer et de les lier à la question du rapport entre la question des " niveaux " de la vie. Pourtant, savoir si ces glissements conceptuels à l'égard de la théorie de l'évolution entraînent, par une sorte d'effet de recul indirect, des modifications – même mineures – sur la question du normal et du pathologique, voilà une question qui devra être laissée pour une étude ultérieure et, sans doute, plus générale.

Nous allons procéder en deux pas. D’abord, il s’agira d’indiquer, sans nécessairement l’expliciter, la différence entre les deux normativités biologiques, celle qui est impliquée dans le rapport entre santé et maladie et celle qui doit être liée aux considérations de Canguilhem sur la théorie de l’évolution. Deuxièmement, nous allons exposer les glissements, les modifications – peut-être moins de substance que sans doute d’accent – qui sont sensibles dans les trois textes successifs mentionnés plus haut à propos de la théorie de l’évolution.

2. La différence entre les deux normativités biologiques

Distinguer entre deux sens de la normativité biologiques est, au premier regard en tout cas, une chose assez facile à partir des textes de Canguilhem. Ainsi, nous citons l’*Essai* de 1943 : “ L’anomalie éclate dans la multiplicité spatiale, la maladie éclate dans la succession chronologique. Le propre de la maladie c’est de venir interrompre un cours, d’être proprement critique. Même quand la maladie devient chronique, après avoir été critique, il y a un autrefois dont le patient ou l’entourage garde la nostalgie. On est donc malade non seulement par référence aux autres, mais par rapport à soi ” (NP, 86-87). De l’autre côté, revient Canguilhem, “ Le porteur d’une anomalie ne peut donc être comparé à lui-même ” (NP, 87). C’est dans cette distinction entre le chronologique et le spatial que loge la distinction entre deux normativités biologiques : d’un côté, la maladie implique ce sentiment, proprement subjectif, de passage, de transition d’une vie à une autre, ce type de limitation ou de dégradation des possibilités de vie à laquelle la maladie nous soumet. D’un autre côté, pour ce qui est de la sélection naturelle – à savoir du propre de la théorie de l’évolution d’inspiration darwinienne –, le porteur d’une anomalie ne peut pas, selon les mots de Canguilhem, et, plus encore – surtout s’il faut se tenir à une vision orthodoxe de la théorie de la sélection naturelle –, *ne doit* même pas être comparé à lui-même. Ce à quoi il faut le comparer, ce sont les autres individus de la même population, les autres individus qui sont d’une même espèce mais qui sont porteurs d’un trait différent ou bien d’une valeur différente d’un même trait. La maladie, autrement dit, implique une *différence* chronologique – une transition intra-individuelle, un passage d’un type de vie à un autre pour un même individu ; de l’autre côté, la sélection naturelle implique une différence spatiale à l’égard d’un même trait, une différence interindividuelle sans laquelle aucune sélection naturelle ne peut être indiquée comme agissant dans une population donnée.

Cette distinction entre une variation intra-individuelle et une variation interindividuelle est reprise explicitement, à propos de la seule maladie et pour renforcer la différence intra-individuelle que celle-ci implique, quelques années plus tard dans l'article sur " Le normal et le pathologique " : " Si ce qui est normal ici, peut être pathologique là, il est tentant de conclure qu'il n'y a pas de frontière entre le normal et le pathologique. D'accord, si l'on veut dire que d'un individu à l'autre la relativité du normal est la règle. Mais cela ne veut pas dire que pour un individu donné la distinction n'est pas absolue " (CV, 207-208). D'un autre côté, la notion de variation interindividuelle est elle-même renforcée dans ce texte par le biais d'une nouvelle définition de l'anomalie, nous citons : " Une anomalie c'est étymologiquement une inégalité, une différence de niveau " (CV, 200). C'est, toujours en suivant une notion orthodoxe de sélection naturelle, cette différence de niveau par rapport à un trait précis ou aux valeurs de ce trait qui est responsable pour l'action de la sélection naturelle.

Enfin, et pour finir ce bref compte rendu de la distinction entre deux sens au moins possibles sinon actuels de la notion de normativité biologique, " Les nouvelles réflexions " de 1966 reprennent un travail de John A. Ryle pour dire : " Ryle distingue enfin entre deux sortes de variations relativement à la norme, à propos desquelles il peut se faire que l'on ait à décider de l'anormalité, en vue de certaines résolutions à prendre d'ordre pratique : variations affectant un même individu selon le temps, variations, à un moment donné, d'un individu à l'autre dans une espèce " (NP, 202). Et Canguilhem continue, avec une phrase qui ouvre toute notre démarche ici : " Ces deux sortes de variations sont essentielles à la survivance " (NP, 202).

Comme nous l'avons déjà annoncé, nous n'allons pas nous occuper ici du rapport direct entre ces deux manières de poser le problème de la normativité biologique, nous n'allons pas essayer de voir directement de quelle manière cette notion de normativité biologique est appelée à répondre à la fois aux exigences d'une variation temporelle intra-individuelle et d'une variation essentiellement inter-individuelle comme celle qui est intrinsèquement exigée par la théorie de l'évolution par sélection naturelle. Au contraire, nous allons nous contenter de suivre ou d'esquisser les différentes manières dont Canguilhem envisage cette dernière, à savoir la théorie de l'évolution. Nous allons donc passer directement à l'analyse des trois manières canguilhemiennes de l'envisager, en suivant l'ordre chronologique des trois textes portant sur le normal et le pathologique.

3. La non-indifférence de la vie dans l'*Essai* de 1943

C'est par rapport à la non-indifférence de la vie qu'il faut comprendre le rapport de Canguilhem à la théorie de la sélection naturelle dans l'*Essai* de 1943. Cela est indiqué par l'importance proprement stratégique que la discussion de la sélection naturelle a dans le discours de l'*Essai*. Canguilhem vient d'affirmer que la mécanique moderne, en tant qu'elle est fondée sur le principe d'inertie, rend absurde la distinction entre les mouvements naturelles et les mouvements violents, car l'inertie est " précisément l'indifférence à l'égard des directions et des variations du mouvement " (NP, 79). Mais, ajoute Canguilhem tout de suite, " la vie est bien loin d'une telle indifférence à l'égard des conditions qui lui sont faites, la vie est polarité ", et il conclut son paragraphe d'une manière nette et dépourvue de prétention : " Tel est le simple fait que nous voulons désigner en parlant d'une normativité biologique " (NP, 79). C'est en ce point précis, pour appuyer cette idée que la vie n'est pas indifférente à ses propres conditions que Canguilhem introduit sa discussion de la sélection naturelle, car, sans aucune transition, le paragraphe suivant commence ainsi : " Il y a des esprits que l'horreur du finalisme conduit à rejeter même la notion darwinienne de sélection par le milieu et la lutte pour l'existence " (NP, 79). Or, à ces esprits, Canguilhem n'oppose pas un finalisme fort, mais un finalisme faible, ou plutôt un finalisme " en négatif " avec cette thèse en quelque sorte minimaliste que la vie, sans se donner une direction quelconque, n'est pourtant pas séparable d'une non-indifférence essentielle à ses propres conditions. Le minimalisme et le caractère à la fois constatatif et non prétentieux est souligné dans la suite du paragraphe lorsque, renvoyant à Georges Teissier, Canguilhem remarque : " parce que beaucoup d'êtres meurent avant que leurs inégalités les servent, cela n'entraîne pas que présenter des inégalités soit biologiquement indifférent. C'est précisément le seul fait dont nous demandons qu'il nous soit accordé. Il n'y a pas d'indifférence biologique. Dès lors, on peut parler de normativité biologique " (NP, 79).

Avec cette thèse d'une non-indifférence biologique dans la théorie de l'évolution, il semblerait qu'on retrouve et généralise le diction que Heidegger récitait à propos du *Dasein* et donc que Canguilhem dirait, en quelque sorte, qu'un vivant est un être pour lequel il y va, en son être, de cet être. Pourtant, il n'en est rien, car ce n'est pas au niveau de l'individu que cette question se pose, mais bien au niveau inter-individuel, comme nous l'avons indiqué plus haut. Ce n'est pas un individu qui " porte " des inégalités, car les inégalités sont par définition relatives : c'est *entre* des individus que des inégalités existent. Il ne s'agit donc pas, dans cette discussion de la non-indifférence, de dire que les individus ne

sont pas indifférents à leurs conditions de vie – ce versant de la discussion appartenant plutôt à l’autre normativité biologique, celle qui renvoie au rapport santé-maladie –, mais plus neutralement ou moins subjectivement, il s’agit de dire que c’est *la vie* qui est non-indifférente à ces inégalités entre les individus. C’est une différence de niveau qui s’opère ici.

En d’autres termes et plus brutalement, ce qui est désigné dans ce point précis par Canguilhem comme non-indifférence de la vie est *le fait même* de la sélection naturelle, ou bien le simple fait *qu’il y a* de la sélection naturelle. Voici un point qu’on pourrait exprimer à partir de Darwin lui-même, avec une citation de l’*Origine des espèces* :

“L’homme ne peut agir que sur les caractères extérieurs et visibles. La nature (...) ne s’occupe aucunement des apparences, à moins que l’apparence n’ait quelque utilité pour les êtres vivants. La nature peut agir sur tous les organes intérieurs, sur la moindre différence d’organisation, sur le mécanisme vital tout entier. L’homme n’a qu’un but : choisir en vue de son propre avantage ; la nature, au contraire, choisit pour l’avantage de l’être lui-même.”⁹

Mais là aussi, à la toute fin de ce passage, Darwin dévie un peu de son propos, car ce n’est évidemment pas pour l’avantage de “l’être lui-même” que la nature choisit, mais bien pour son avantage à elle. C’est peut-être pour cela que Darwin revient avec l’affirmation suivante :

“On peut dire, par métaphore, que la sélection naturelle recherche (*scrutinises*), à chaque instant et dans le monde entier, les variations les plus légères.”¹⁰

Voici une formule plus juste, car la non-indifférence de la vie – ou de la nature, selon les mots de Darwin – n’est rien de plus que cette “recherche” – mal traduite à partir de *scrutinises* –, ce regard attentif à tous les détails, qui n’a évidemment pas de direction préétablie – ce que Canguilhem souligne lui-même en disant que milieu et vivant ne sont “normaux” que dans leur rapport – mais qui, justement, ne reste pas et ne peut pas rester indifférent aux différences entre des individus. C’est une coextensivité de la vie et de la sélection que Darwin énonce ici et que Canguilhem reprend sous le nom de non-indifférence de la vie : c’est l’idée qu’en biologie évolutionniste être signifie être sélectionné, ou bien, plus précisément, qu’en biologie évolutionniste *vivre signifie essentiellement continuer à vivre*.

⁹ Charles Darwin, *De l’origine des espèces*, Paris, GF Flammarion, 1992, p. 132.

¹⁰ *Ibid.*, p. 133.

Sans doute, quelques pages plus loin dans l'*Essai*, revenant à la question de la sélection naturelle, Canguilhem essaie d'aller un peu plus loin en insistant sur les vertus de la variabilité des vivants d'une espèce contre une spécialisation rigide. Mais cette discussion reste encore assez hésitante et un peu confuse car elle reste, d'une part, sous la tutelle de l'idée générale et neutre de non-indifférence biologique telle que nous l'avons exposée ; et, d'autre part, cette discussion reste sous la tutelle, dérivée de la première, de la " normalité " purement relationnelle du milieu et des espèces vivantes, dans laquelle pourtant la notion de milieu reste quelque peu restrictive, insuffisante, car, nous allons le voir, trop empreinte de mécanisme . Voici une citation qui peut éclairer ce point :

"Le milieu est normal pour une forme vivante donnée dans la mesure où il lui permet une telle fécondité, et corrélativement une telle variété de formes, que, le cas échéant de modifications du milieu, la vie puisse trouver dans l'une de ces formes la solution au problème d'adaptation qu'elle est brutalement sommée de résoudre. Un vivant est normal dans un milieu donné pour autant qu'il est la solution morphologique et fonctionnelle trouvée par la vie pour répondre à toutes les exigences du milieu." (NP, 90-91)

Tant que le rapport entre milieu et vivants est gardé dans une symétrie stricte entre quelque chose de biologique et quelque chose de mécanique – et c'est un point que Canguilhem tentera de changer plus tard –, il devient difficile de comprendre quel est l'agent ici, il devient difficile de comprendre comment une variation inutile dans un milieu donné actuellement peut apparaître en vue des milieux éventuels. A lire la première phrase citée ci-dessus, on dirait que le vrai agent dans cette création de variabilité est le milieu lui-même car, aux dires de Canguilhem, il ne sera dit normal que s'il " permet " – et non s'il occasionne ou même rend possible – à la forme vivante donnée un tel degré de variabilité. C'est sur ce rapport du milieu et des formes vivantes que Canguilhem reviendra pour prolonger ses vues sur la sélection, mais dans l'*Essai* il ne peut que trancher en revenant ou en insistant sur la non-indifférence de la vie comme simple fait de la sélection naturelle : " Relativement à toute autre forme dont il s'écarte, un vivant est normal, même s'il est relativement rare, du fait qu'il est, par rapport à elle *normatif*, c'est-à-dire qu'il la dévalorise avant de l'éliminer " (NP, 91). C'est à l'intérieur d'une même espèce, entre des types – ou bien, selon les mots de Teissier, des " races " différentes –, que la normativité biologique conjuguée comme sélection naturelle s'exprime. Canguilhem ne recule donc pas devant les résonances guerrières ou brutales de la sélection naturelle, comme le montre aussi le passage suivant : " Le normal, en matière biologique, ce n'est pas tant la forme ancienne

que la forme nouvelle, si elle trouve les conditions d'existence dans lesquelles elle paraîtra normative, c'est-à-dire déclassant toutes les formes passées, dépassées et peut-être bientôt trépassées ” (NP, 91). Voici donc ce qui constitue, à notre avis, la première étape du rapport de Canguilhem à la sélection naturelle : la non-indifférence de la vie, dans l'*Essai* de 1943, peut en toute vraisemblance être resserrée autour du fait de la sélection naturelle, ou bien du fait qu'il y a de la sélection naturelle.

4. La hiérarchisation des formes de vie possibles dans l'article sur “ Le normal et le pathologique ”

L'accent se déplace pourtant, huit ans plus tard, dans l'article sur “ Le normal et le pathologique ” publié par Canguilhem en 1951. Une citation de cet article nous montrera l'ampleur de ce déplacement – nous citons : “ dans l'ordre de la vie, il n'y a pas de réussites qui dévalorisent radicalement d'autres essais en les faisant apparaître manqués ” (CV, 201). Au premier regard, rien ne semble radicalement changé par rapport à l'*Essai* de 1943, car une citation de cet essai que nous avons donnée plus haut nous disait qu'une forme vivante est normative seulement *relativement* à une autre dans la mesure où elle la dévalorise avant de l'éliminer. Ici, nous semblons avoir affaire à une affirmation similaire, prise du côté absolu et non du côté du “ relatif ”, dans le sens où Canguilhem nous dit qu'une forme de vie ne peut en dévaloriser *radicalement* une autre, ou, comme il l'avait dit quelques lignes auparavant, nous citons, “ il ne peut rien manquer à un vivant ” ou bien, plus clairement, “ il n'y a pas en soi et à priori de différence entre une forme réussie et une forme manquée ” (CV, 201).

Pourtant il y a un changement ici, et il est introduit par le fait que Canguilhem ne parle plus seulement de formes vivantes, mais bien d' “ essais ” : il n'y a pas, nous dit-il, de réussites qui dévalorisent radicalement d'autres *essais*. Ce point marque une modification significative par rapport à l'*Essai* de 1943 dans la mesure où la vie n'est plus seulement non indifférente, mais elle est aussi “ *aventureuse* ”. Le non-indifférent, c'est celui qui ne veut pas sortir en perte d'une situation donnée, alors que l'aventureux c'est celui pour lequel la question de la perte ne se pose plus, ou plutôt celui qui se comporte comme si, de tout façon, dans n'importe quelles circonstances actuelles ou éventuelles, un gain ne pourrait pas ne pas se présenter, sous n'importe quelle forme. C'est dans le sens de cet esprit aventureux que Canguilhem interprète ici les singularités individuelles en disant

qu'on peut les interpréter " comme un échec ou comme un essai, comme une faute ou comme une aventure " (CV, 200). Comme un échec, c'est-à-dire comme un écart inexplicable par rapport aux lois de la nature conçues comme des " essences génériques éternelles " dont la non-application exacte serait due, selon une superbe formule, à une nature " trop riche pour se conformer à sa propre économie " (CV, 199). Mais comme un essai ou comme une aventure, de l'autre côté, dans le sens où, nous dit Canguilhem, il s'agit " des organisations dont la validité, c'est-à-dire la valeur, est référée à leur réussite de vie éventuelle " (CV, 200). Or, si c'est l'avenir qui décide de la valeur des formes vivantes, alors, conclut Canguilhem en une phrase très forte à laquelle il nous faudra revenir, " les réussites sont des échecs retardés, les échecs sont des réussites avortées " (CV, 201).

Il faut pourtant s'arrêter un peu ici pour bien comprendre ce que Canguilhem nous dit, en effet. Si les formes vivantes et les singularités individuelles sont des essais ou des aventures, ce ne sont pourtant pas eux qui sont aventureux. Ce qui, en eux, est aventureux, ce dont ils manifestent le caractère aventureux, ce sont bien les espèces qui produisent des types d'individus variables ou des singularités individuelles. C'est, autrement dit, à un autre changement de niveau, différent de celui de l'*Essai*, que nous avons affaire ici, comme l'indique la phrase suivante : " Un genre vivant ne nous paraît pourtant un genre viable que dans la mesure où il se révèle fécond, c'est-à-dire producteur de nouveautés, si imperceptibles soient-elles à première vue " (CV, 199). Ce ne sont plus ou plus seulement les formes vivantes – c'est-à-dire les types d'individus d'une même espèce – qui doivent être jugées viables ou non, mais ce sont les *genres* vivants, c'est-à-dire les espèces elles-mêmes. La normativité biologique qui est à l'œuvre ici n'est pas seulement la non-indifférence de la vie, c'est-à-dire le fait que, *à l'intérieur d'une même espèce*, il y a de la sélection naturelle entre des formes ou des types d'individus différents. Ce n'est plus seulement au simple fait de l'existence de la sélection naturelle qu'on doit assigner la normativité biologique, mais aussi à la production de nouveautés ou à la capacité de créer de la variabilité. Un genre ou une espèce est plus viable qu'un autre dans la mesure où il est capable de produire plus de variabilité, et ainsi de faire face non seulement au milieu actuel, mais aussi à des milieux éventuels ou à des changements éventuels du milieu actuel.

Sans doute, cette modification d'accent qui s'opère entre l'*Essai* et l'article de 1951 est due à une modification ou plutôt à un approfondissement de la notion de milieu, un approfondissement auquel Canguilhem s'était consacré en 1946-1947 et qui a donné lieu à la conférence reprise dans *La conscience de la vie* sous le titre " Le vivant et le milieu ". Le milieu n'y est plus vu comme une sorte de cadre inerte ou mécanique – tel qu'il est, par la

neutralité même de l'usage du terme, employé dans l'*Essai* – mais il devient le corrélat, lui-même intérieurement différencié ou non homogène, d'un comportement. Dans ce sens, Canguilhem ne repousse pas, au contraire, la notion d'adaptation au milieu, mais la complique en indiquant qu'un même "environnement" peut être le milieu différentiel pour plusieurs manières de vivre ou de se comporter dans le cadre d'une même espèce. C'est ce que, plus tard, le petit livre cosigné par Canguilhem *Du développement à l'évolution au XIX^e siècle*, affirmera le plus clairement dans une phrase qui sera reprise par Canguilhem dans les "Nouvelles réflexions sur le normal et le pathologique" : "Les places vacantes en un lieu donné, selon la terminologie de Darwin, sont moins des espaces libres que des systèmes de vie (habitat, mode d'alimentation, d'attaque, de protection) qui y sont théoriquement possibles et non encore pratiqués".¹¹ Si une espèce peut être jugée viable en fonction de sa capacité de créer de la variabilité, elle peut l'être seulement en relation avec un tel environnement qui n'est lui-même pas homogène ou qui est intérieurement différencié selon les opportunités d'attaque ou de protection, d'alimentation, etc.

Mais alors, pour revenir à notre question, si Canguilhem peut dire que la valeur d'une forme de vie – c'est-à-dire d'un type d'individu d'une espèce – est référée à sa réussite de vie éventuelle, il faut bien comprendre que cette réussite de vie appartient autant à ce type de vie et à son espèce. La réussite éventuelle d'un type n'est pas seulement sa réussite à lui, mais tout autant la réussite de l'espèce qui a créé assez de variabilité pour qu'une telle réussite puisse avoir lieu. C'est ici que s'opère, à notre avis, la modification la plus importante par rapport à la non-indifférence de la vie de l'*Essai*, dans la mesure où la vie n'est plus ou plus seulement définie comme non-indifférence, mais d'une manière plus profonde comme "tentative de hiérarchisation des formes possibles" (CV, 200-201). La vie n'est pas seulement non-indifférente à ses conditions actuelles, elle relève aussi d'un rapport essentiel au possible. Mais ce rapport au possible ne peut pas se comprendre sans un rapport intrinsèque avec la hiérarchisation, ou plutôt, comme nous avons essayé de le dire, avec cette idée que la réussite ne peut pas appartenir à un seul niveau de l'organisation biologique, mais se répand immédiatement, automatiquement, on pourrait dire, sur la totalité de l'organisation biologique. Le succès, la survie ou la multiplication, ce n'est pas seulement la réussite du type d'individus qui l'atteint, mais aussi le succès de son espèce.

¹¹ G. Canguilhem, G. Lapassade, J. Piquemal, J. Ulmann, *Du développement à l'évolution au XIX^e siècle*, dans *Thalès*, XI, 1960, p. 32.

Voici un point essentiel qui, nous semble-t-il, n'est pas inexistant dans la biologie contemporaine. C'est ainsi que, par exemple, Stephen Jay Gould écrit dans *The Structure of Evolutionary Theory*, qu'une certaine " intuition " s'est installée parmi les chercheurs de la question de la sélection naturelle à des niveaux multiples, une intuition conformément à laquelle la vraie sélection au niveau de l'espèce est celle qui donne lieu à des taux de spéciation différents, tandis que les taux d'extinction différents sont réductibles à des processus relevant du niveau de l'organisme. Et, dit Gould, " la source de cette intuition – qui pourrait s'avérer erronée ou bien superficiellement ancrée – vient d'une sorte de saisie du fait que l'extinction d'une espèce peut souvent être adéquatement expliquée simplement comme l'addition des morts de tous les organismes de l'espèce, chaque mort survenant pour des raisons concernant l'organisme seul et sans la contribution des propriétés appartenant au niveau de l'espèce. Lorsque le dernier organisme reproductif meurt, l'espèce est disparue. Mais comment pourrait *surgir* une nouvelle espèce sans quelque implication de la part des traits qui se situent au niveau des populations ? Après tout, les organismes ne sont pas sujets à la spéciation ; seulement les populations le sont ".¹²

Il est peu important de remarquer ici que Gould parle de l'apparition des nouvelles espèces alors que Canguilhem parle de leur modification dans des milieux éventuels (pouvant, tout de même, mener à de nouvelles espèces). Mais ce qui compte, c'est ce point étrange que la mort, ou bien l'échec, appartient aux individus, alors que la réussite appartient toujours à une pluralité de niveaux de l'organisation biologique. C'est pour cela que Canguilhem parle des réussites comme d' " échecs retardés " et des échecs comme de " réussites avortées ", car dans ces expressions il y va seulement de la réussite ou bien des échecs des formes de vie, c'est-à-dire des types déterminés d'individus. Or, en changeant de niveau, pour parler, non plus seulement des formes de vie, mais pour y intégrer aussi les " genres " de vie, nous approchons ici un point essentiel qui pourrait être celui de la définition de la vie comme réussite, et comme réussite toujours hiérarchisée. Il ne s'agit plus de déterminer une essence de la vie, de la localiser dans un principe ou dans une sorte de force aveugle, il s'agit moins de parler d'une créativité de la vie, d'une prodigalité ou d'un gaspillage qui la caractériserait, que de voir dans la vie une sorte de réussite intermittente qui, en tant que réussite, ramasse, pour ainsi dire, à chaque fois, tous ses niveaux. La vie n'est ainsi ni force ni création, elle est ce point ou cette série intermittente

¹² Gould, *The Structure of Evolutionary Theory*, Cambridge, Belknap Press, 2002, p. 669 (la traduction nous appartient).

d'émergence où son prolongement devient inséparable de sa réussite étalée sur tous ses niveaux. Pour transformer une phrase de Canguilhem où il dit qu'un homme est normal seulement lorsqu'il est " plus que normal " (CV, 210), nous dirions qu'en ce point " vivre " et " la vie " ne se superposent plus d'une manière nette, qu'en ce point, pour ainsi dire, la vie n'est vie qu'en étant plus que " vivre ".

A la non-indifférence de la vie que présentait l'*Essai* il faut donc ajouter, avec l'article de 1951, l'idée de la vie comme réussite hiérarchique. Mais, dans l'état de la discussion qui est celle de l'article de 1951, il ne faut pas trop voir dans cette notion de réussite hiérarchique, il ne faut pas la transformer dans une position forte et elle-même rigide sur l'évolution. Car, si Canguilhem peut dire que c'est seulement l'avenir des formes vivantes qui décide de leur valeur, on peut sans doute dire la même chose non seulement des formes vivantes, mais aussi des " genres vivants ", c'est-à-dire des espèces. Canguilhem reste ici prudent, et il n'envisage pas de statuer d'une manière tranchée sur le point de savoir si on pourrait juger ou si on pourrait se prononcer sur la valeur d'une espèce à partir de sa seule capacité de créer de la variabilité. Cette prudence est manifeste, par exemple, dans le fait que Canguilhem mentionne la possibilité – mais la possibilité seulement – de gènes de mutabilité (c'est-à-dire de gènes qui multiplieraient la latitude de mutation d'autres gènes), mais n'en fait pas un grand usage. Et la prudence dont Canguilhem témoigne ici pourrait être résumée en disant que les espèces, en créant de la variabilité, se donnent la possibilité d'un avenir, mais ne se l'assurent pas. La variabilité est seulement l'ouverture d'une possibilité, et non une garantie. C'est pour cela d'ailleurs que Canguilhem définit dans cet article le monde vivant comme une *tentative* de hiérarchisation des formes possibles, et l'accent ici doit être mis sur le mot " tentative ". C'est ce seuil ou cette prudence qui semble avoir été franchi par Canguilhem dans les " Nouvelles réflexions concernant le normal et le pathologique ", vers lesquelles nous nous tournons maintenant.

5. La sélection naturelle dans les " Nouvelles réflexions " de 1966

Voici le passage qui peut introduire ce nouveau pas effectué par Canguilhem : " une théorie mutationniste de la genèse des espèces ne peut définir le normal que comme le temporairement viable. Mais à force de ne considérer les vivants que comme des morts en sursis, on méconnaît l'orientation adaptative de l'ensemble des vivants considérés dans la continuité de la vie, on sous-estime cet aspect de l'évolution qu'est la variation des modes

de vie pour l'occupation de toutes les places vacantes ” (NP, p. 198). A un premier regard, rien ne semble changé, mais il faut que nous soyons tout de même attentifs au fait qu'il s'agit, ici, d'une reprise des deux positions antérieures de Canguilhem. D'un côté, il s'agit de reprendre une thèse de Teissier contre les mutationnistes¹³ pour dire que présenter des indifférences biologiques n'est pas indifférent pour les vivants : le “ temporairement viable ” ici signifie en effet ce qui a été, selon l'image du mutationnisme donnée par Teissier, laissé subsister par le hasard seul. Pourtant, la deuxième phrase de cette citation va plus loin, en énonçant non seulement que les vivants ne sont pas des morts en sursis parce que leurs traits ne sont pas indifférents à leur survie, mais également parce que la création de variabilité qui est à l'œuvre en eux contribue elle aussi – à côté de leurs traits, pour ainsi dire – à leur succès. Et cela est d'autant plus significatif que le rejet de l'image des “ morts en sursis ” va en quelque sorte contre l'idée même de l'article de 1951 que les réussites sont des échecs retardés.

C'est cette combinaison des deux arguments à la fois qui, il nous semble, marque la nouveauté, au moins partielle, des “ Nouvelles réflexions ” sur la question de la sélection naturelle. Elle est indiquée par l'usage que Canguilhem fait de la notion d' “ homéostasie génétique ” de Michael Lerner, à savoir de la tendance des populations biologiques à garder leur composition génétique. D'une manière incidente et ironique, Lerner lui-même tient ce terme pour équivalent de celui “ d'inertie génétique ”. Mais cela est moins important, ce qui est plus significatif, c'est que, en suivant dans une certaine mesure Lerner lui-même, Canguilhem considère, comme principal facteur responsable de cette homéostasie, la supériorité des hétérozygotes. Qui plus est, il tient cette supériorité des hétérozygotes comme étant liée à l'hétérosis, à savoir à la viabilité et à la fécondité supérieures des hybrides dans les animaux et plantes domestiques. Pourtant, des problèmes apparaissent ici. D'abord, dans la supériorité des hétérozygotes que Canguilhem convoque, la référence au milieu semble avoir disparu : s'il cite l'anémie de Cooley comme un cas où l'hétérozygote est supérieur à la fois à l'homozygote mutant et à l'homozygote normal, il ne mentionne pourtant pas le fait que les hétérozygotes pour les trois gènes qui déterminent les deux types d'anémie de Cooley ne sont supérieurs *que* dans les milieux affectés par le paludisme et pas ailleurs, et son affirmation tranchante et non relativisée par le milieu – “ Les hétérozygotes sont plus féconds ” – tend à obscurcir ce fait (NP, p. 199). En outre, même si dans les années 1950, la supériorité des hétérozygotes semblait être la meilleure explication pour la viabilité

¹³ Cf. Georges Teissier, “ Mécanisme de l'évolution I », in *La Pensée*, 1945, no. 1, pp. 14-15.

et la fécondité accrues des hybrides, des expériences ultérieures ont montré¹⁴ que c'est sont les phénomènes de dominance – le fait que l'expression des allèles récessifs désavantageux est plus entravée dans les hybrides – et non pas ceux de surdominance qui expliquent cette supériorité des hybrides.

Sans doute, par rapport à ce dernier point, pour reprendre une phrase de Canguilhem, on ne peut pas lui reprocher de ne pas connaître des choses qu'il ne pouvait pas connaître. Mais ce qui est important, c'est que dans les “ Nouvelles réflexions ” Canguilhem semble dire non seulement qu'une variabilité accrue peut se révéler importante pour la réussite d'une espèce ou d'un type, mais, semble-t-il, qu'une variabilité accrue est supérieure à une variabilité moindre d'une manière absolue. Dans n'importe quelle situation, la variabilité accrue est “ plus forte ”, pour ainsi dire, voilà ce qui semble soutenir, d'une manière allusive, sans doute, l'importance que Canguilhem accorde ici à la supériorité des hétérozygotes. La vie n'est plus seulement non indifférente par rapport à ses conditions actuelles, elle n'est plus seulement l'ouverture des possibilités de répondre à des problèmes possibles – c'est-à-dire *tentative* d'une hiérarchisation des formes de vie possibles – mais elle semble devenir un *garantie actuelle* pour toute hiérarchisation éventuelle. Voici l'accent, trop triomphaliste à notre avis, introduit par les “ Nouvelles réflexions ” et paraphé, pour ainsi dire, par une citation que Canguilhem reprend à André Lwoff – nous citons : “ dans la nature il n'y a pas de problèmes ; il n'y a que des solutions ” (NP, p. 199).

Par une sorte de mouvement en harmoniques, les positions de Canguilhem sur la sélection naturelle semblent donc s'élargir dans un premier temps seulement pour se resserrer ensuite. Il semble partir d'une vision monolithique de la non-indifférence de la vie, à savoir du simple fait de la sélection naturelle, sans que cette “ normativité biologique ” trouve un ancrage hiérarchique. Un deuxième pas semble porter Canguilhem vers une considération directe du fait qu'une telle normativité ne peut pas être séparée d'une sorte de retrospectivité essentielle de la vie qui, lors de ses réussites, ne peut pas ne pas ramasser et étaler tous ses niveaux. Le point final pourtant de cette évolution semble mener vers un nouveau rétrécissement dans le sens où les réussites de la vie ne sont plus nécessairement des événements, mais paraissent plutôt devenir des résultats seulement, des dérivations à partir de certains mécanismes simples, comme dans le cas de la supériorité des hétérozygotes, quasiment préétablis comme des solutions antérieurs à leurs problèmes. C'est cette oscillation de Canguilhem autour de la question de l'évolution et de la hiérarchie

¹⁴ Cf. James Crow, “ 90 Years Ago : The Beginning of Hybrid Maize ”, *Genetics* 148, pp. 923-928.

que nous avons voulu présenter ici sans pourtant prétendre en pouvoir tirer des conclusions fortes qui, quant à elles, ne pourront pas être tirées sans une étude compréhensive de l'œuvre entière de Canguilhem. C'est d'ailleurs pour cela que nous laissons ouverte la possibilité que cette oscillation ait une signification pour la question, plus propre à Canguilhem, de la normativité biologique impliquée dans le rapport entre santé et maladie.

Is a naturalistic concept of health possible?

Canguilhem's and Boorse's Theories of Normality*

Some years ago as I played with my younger son, I felt pain in my back and we had to abandon our little amusement. After two days my kid asked me full of expectation: "Can we play again, did you repair your back?" The almost same meaning of 'repairing' and 'healing' seems to be the part of our language intuition which supports the familiar philosophical analogy between organisms and artifacts. My analysis of the concept of health will start with that alleged similarity of functional language we use when we speak of organism's parts and devices and with the question what it means to ascribe normal functioning to something. Namely, one of the most important ideas of the contemporary naturalistic conception of health is precisely the idea of biological function and its connection with normality. Moreover the possibility of an objective, value-free concept of health is usually associated with the notion of function. It takes little to see how the question 'what is function?' is connected to the question 'what is health?' We often describe 'disease', 'deformation' and 'disability' as 'dysfunction'. However there are many conceptions of biological function and each of them has its implications for interpretation of normality. The two most relevant are the proper function conception and the causal role function conception (CR). The proponents of the first consider the function in context of evolution, while their opponents claim that the function of an organism's part has nothing to do with the evolutionary history. It means that functional language in science is based on quite opposite theories of biological function. I argue that some of the most important theories of health can be seen as examples of such kind.

In the first section I shall explore the conceptions of biological function and their importance for the understanding of health. The aim of this section is to show how the particular notion of *proper function* generates a "quasi-normative...notion 'Normal'"

* This paper was supported by the projects No. 179041 (Dynamic Systems in Nature and Society: Philosophical and Empirical Aspects) and No. 41004 (Bioethical Aspects: Morally Acceptable Within the Biotechnologically and Socially Possible) financed by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

which plays a fundamental role in the biological and medical conception of health¹, and whether that notion can really help us to define health and disease. In the second section it will be shown that some of the major difficulties in influential conceptions of health like Boorse's stem from mixing two notions of function in definition of health: the proper function and the CR function. The former is needed to introduce the element of species design in the definition of health, and the latter to define the concept of health as a functional contribution to the physiological goals of organism. I intend to show further that the concept of CR function is inappropriate to serve as a foundation for the concepts of health and disease. Yet, the aim of this article is not primarily to criticize Boorse's naturalistic view of health, but much more to analyze the relationship between function and norm. Namely, no statement of biological function is really free from norm-component. Canguilhem's view which is constructed in opposition to physiological notions of the normal and the pathological is based on a different understanding of the normativity of organisms. I hope that some intuitions which can be found in Canguilhem can help us to come to a concept of health which is naturalistic and incorporates normativity at the same time.

1. Where do biological norms come from?

It is clear that functional analysis can be applied to very different fields of reality: its subject may be a function of an organ, but also a function of an artifact, or a conscious activity. Already in the area of artifacts, where it seems that things are pretty clear, functional analysis is faced with many difficulties. We can, for example, wonder whether the function of an object is determined by the intentions of the person who made it or by the intentions of the person who uses it: the hammer is made for hammering nails, but it can also serve to press paper. Also, it is imaginable that the intentions of the designer completely differ from the actual use of his artifact, so the intentions could be irrelevant for functional analysis. Further, and that is of particular importance in our context, we can speak about a function of an artifact which has never been used, and we can also speak about a function of an artifact independently of the way in which it is actually being used,

¹ Ruth Garrett Millikan, *Language, Thought, and Other Biological Categories*, Cambridge Massachusetts, MIT Press, Bradford Book, 2001, p. 5.

and even, irrespective of the question whether it is properly functioning or not. The function of the machine for baking bread is baking bread regardless of the fact that the machine may be faulty. And what is happening with the organisms? Is it really justified to interpret organisms like artifacts? I do not claim that the similarity we see between organisms and artifacts is part of some original human intuition. It is possible that this kind of intuition exists, but it is not a necessary assumption for my analysis, because it takes its point of departure in the use of functional language in biology. Many biological investigations examine the function of an organism's trait, but certain ambiguities, which call into question the legitimacy of functional language, are peculiar to this practice. Here I will analyze some of them which are important for our understanding of health and disease, particularly the incompatibility of two function conceptions, the so called proper function and the causal role function.

According to the proper function theory, functional explanations must contain the statements about the origin of the function carriers, that is, functional explanations “concern the causal background of the phenomena under consideration.”² The functions were determined in terms of selective effects: “biological proper functions are effects for which traits were selected by natural selection”³. Therefore, vertebrates do not have the heart because of their actual effect, blood pumping, but because the heart was selected in the ancestral population as an adaptive trait, a trait which helps the organism to adapt to its environment. It means that the functional ascriptions are statements about evolution of the population, rather about how a trait (for example the heart) has come to be there. The significant component of the theory of proper function is the descriptive normativity, a normativity which “involve some kind of non-arbitrary framework allowing comparison between actual and alternative states”⁴. Yet another thing is also important for a proper function: namely, that conception is to a large degree connected with the idea of design. I would like to explain briefly these two ideas, design and descriptive normativity, and their relevance for the notion of health. As Neander says, “...the normativity of biological functions is neither simply evaluative nor statistical.”⁵

² Larry Wright, “Functions”, *The Philosophical Review*, vol. 82, n° 2, 1973, p.156.

³ Karen Neander, “Functions as Selected Effects: The Conceptual Analyst’s Defense”, *Philosophy of Science*, vol. 58, n° 2, 1991, p.168.

⁴ Wayne Christensen, “Natural Sources of Normativity”, *Studies in History and Philosophy of Biological and Biomedical Sciences*, n° 43, 2012, p. 105.

⁵ Karen Neander, “Misrepresenting and Malfunctioning”, *Philosophical Studies*, n° 79, 1995, p. 111.

The notion of proper function is normative: it means that, among other things, we can speak about malfunctioning. We can speak about ‘normal’, ‘standard’ functioning and about ‘abnormal’ functioning as a deviation from the standard. The functional statements imply the norm which makes it possible to talk about dysfunctions, because the organs or processes were designed to perform some proper functions, regardless of whether they are now able to perform them or not. ‘Biological norm’ has here a broad meaning of a standard from which it is possible to deviate. What’s it all about? Actually, functional analysis of some organism’s component includes not only its current performances, but also the assumptions about what the component is supposed to do. This assumption is the frame of reference in relation to which the current function of an organ will be judged. However, it is not logically primary, because we first have to know why the trait was selected in the past. In order to explain why a trait is now dysfunctional, then, we have to find what the trait is supposed to do, and this requires the knowledge about the natural selection in the past and past fitness of the trait. The functional analysis is a kind of historical analysis⁶, and function is adaptation.

Let us see what the theory of CR function claims. According to that theory, function of heart, for instance, has nothing to do with its presence in the vertebrates. In other words, the presence of the heart can’t be explained “by appeal to what the heart *does*”.⁷ Moreover, function has nothing to do with evolution and adaptation. Why should we have to know that there was natural selection for the heart in the ancestral population to understand what the function of the heart is? It is sufficient to know that thanks to the heart nutrients are transferred through the body. Therefore, to determine the function of a body part it is sufficient to know which contribution the part under consideration makes to the capacities of the body. Namely the fact that organism possesses some part will not be explained by its function but by the information about the inheritance: I have a heart because I inherited it from my parents. It is quite wrong to explain the aforementioned fact by the selective history of the heart. Function and presence of some body part have no immediate connection. An often cited example can help us understand an important difference between two theories. Both fresh-water and marine protozoans incorporate contractile vacuoles. The proponents of proper function would say that protozoans have contractile vacuole because of its function, namely, because it helps the organism to solve the problem

⁶ Ruth Garrett Millikan, “In Defense of Proper Functions”, *Philosophy of Science*, vol. 56, n° 2, 1989, p. 299.

⁷ Robert Cummins, “Functional Analysis”, *The Journal of Philosophy*, vol. 72, n° 20, 1975, p. 745.

of excess water. Because of that effect the organisms with contractile vacuoles have been selected in the evolutionary past. But, on the other hand, the marine protozoans do not have such a problem with water, so the natural selection can't be an explanation of the fact that some organisms incorporate contractile vacuoles. A detailed analysis of functional language in biology shows that proper function account is based on a derived understanding of function, moreover on a parasitic use of language: "...the survival value of a structure hangs on what capacities of the organism, if any, are explicable by appeal to the functioning of s."⁸

Peculiar to CR theory of function is the relationship between the system and its components: we can see not only the organism as a system, we can also analyze the organism's subsystems in terms of their components. The criteria for distinguishing the system from the part are quite liberal, it is relevant only that capacities of the components are simpler than the capacities of the system. This means that the moment of organization is crucial for the functional analysis and that objects to which we can't ascribe complexities in some sense can't be analyzed in functional terms. Moreover, the value of functional analysis depends on the difference in the complexity of the components and of the system. The greater the difference, the more valuable the analysis. But the liberality in relation to the criteria produces some difficulties relevant for our context. Consider an example from pathology. If we are allowed to understand each relatively complex object as a system, then we can apply functional analysis to phenomena like cancer. Strictly speaking, we are obliged to ascribe a positive function to deviant processes in cancer cells (deviant in comparison with healthy cells), because they contribute to the capacities of the cancer as a whole. Namely, thanks to the dysfunctional processes cancer grows and pushes the healthy tissue. On one side, this implies that we can speak about pathological processes in positive terms of contribution to the capacities. It seems that this situation "illustrates an important principle, namely that what is advantageous at one hierarchical level may be disadvantageous at another level, leading to potential conflict."⁹ On the other hand, function as contribution to the capacities of the system lacks any moment of normativity: this functions "are nonetheless purely descriptive (not normative) in the sense that the causal role function of *X* is some subset of *X*'s actual causal role".¹⁰

⁸ *Ibid.*, p. 749.

⁹ Samir Okasha, *Evolution and the Levels of Selection*, Oxford, Clarendon Press, 2006, p. 11.

¹⁰ Neander, "Functions as Selected Effects", p. 181.

Let us pause here to explain a little bit more in detail this particular problem which is related to another relevant aspect of physiological theory of function. The function is here determined by the notion of disposition. In other words, it means that the component to which the function is ascribed must be able to perform the function: “if something functions as a pump in a system *s* or if the function of something in a system *s* is to pump, then it must be capable of pumping in *s*.”¹¹ Or, if something to which we would ascribe function is not capable of performing that function, we can’t speak about function at all. But, the expectation that the component will always really perform the action which is its function is too strong and restrictive, so that referring to ‘disposition’ seems like a way out of those difficulties. We can apply functional analysis only to objects which regularly contribute to capacities of the systems containing them. Those systems “correspond to ideal types, not to actual historical tokens or kinds”.¹² It is not necessary for the components to actually perform the action, it is enough that they have a disposition for performing it. What are the implications for the understanding of health and disease?

We have seen that the contribution to the capacities of the organism is the main criterion for functionality, but how should we understand the diseased heart which does not contribute to these capacities? It seems that the heart that cannot perform its function is no more a heart, and no more within the class of functional objects. A healthy heart is a functional category, but a diseased heart is something completely different. It can no longer be the subject of functional analysis. Why then do we use to talk about diseased organs as dysfunctional organs, what does this mean? Is it a word without meaning or some kind of metaphor? Or must we ascribe a positive function to diseased organs or pathological processes, as we mentioned in the example with cancer? Some authors, like Boorse, tried to save this conception by establishing so called biological goals which would allow them to understand the causal contribution and proper function as the same thing.¹³ Boorse’s position will be discussed more in detail a little later; for now, it is enough to emphasize that the conception of function as contribution is in some relevant respects similar to the older Claude Bernard’s theory of the normal and the pathological. Both conceptions claim that disease, deformation, defect etc. can be determined only negatively, as a lack of

¹¹ Robert Cummins, “Functional Analysis”, p. 757.

¹² Ruth Garrett Millikan, “Biofunctions: Two Paradigms”, in: A. Ariew ed., *Functions*, pp. 113-143. Oxford University Press, 2002, p. 29.

¹³ Christopher Boorse, „What a Theory of Mental Health should be“, *Journal for the Theory of Social Behaviour*, vol. 6, Issue 1, 1976, p. 78.

function, and positive statements about the pathological phenomena are possible merely in terms of quantitative variations of the normal physiological phenomena.¹⁴

I would like briefly to return to the normative aspect of the conception of proper function. According to the proponents of this conception, when we want to explain why a trait is dysfunctional, we have to know what the trait is supposed to do, and to this end it is necessary to know why the trait has been selected in the past. We remember that the current contribution of a trait is irrelevant for the functional analysis: “The relevant consideration is not what an item actually does but what, based on its evolutionary history, it is (was) supposed to do.”¹⁵ Moreover, this conception implies that “the function of kidneys (both normal and abnormal) is to filter wastes from the blood... (and this fact remains true even if renal failure becomes universal).”¹⁶ This means that the functional explanation of some traits depends on their history. One of the interesting consequences is that things which do not have history cannot have a function either. It is possible to imagine an organism, a lion, for example, which comes into being without evolution (through a chance), but is in respect of all its features identical to an African lion which is a product of natural history.¹⁷ Is it justified to ascribe functions to the parts of such an organism? Intuitively, we tend to respond positively to this question. But, the proper function theorist say “No”, because the function is determined by history. “Such a double has no proper functions because its history is not right”.¹⁸ There is no doubt that we can use the notion of function in whatever way we want, but in such a kind of conceptual analysis we are focused on the way in which the scientific community uses the notion.

Historical aspect of the function is related to its normative aspect. That is to say, the norms can be derived from natural history, which also means that “ahistorical theories of functions do not provide a notion of ‘proper function’ which is appropriately normative, and so they cannot accommodate the theoretical role required of the notion in biology”¹⁹. Statements about function are actually statement about past adaptive value of the function bearer. The normative terms which are introduced with the conception of proper function denote a kind of standard for a certain type. So, ‘proper function’ does not mean something

¹⁴ Georges Canguilhem, *The Normal and the Pathological*, New York, Zone Books, 1991, p. 42.

¹⁵ Peter McLaughlin, *What Functions Explain. Functional Explanation and Self-Reproducing Systems*, Cambridge, Cambridge University Press, 2003, p. 107.

¹⁶ Neander, “Functions as Selected Effects”, p. 183.

¹⁷ Christopher Boorse, „What a Theory of Mental Health should be“, p. 74; Neander, “Functions as Selected Effects”, p. 169.

¹⁸ Millikan, „In Defense of Proper Functions“, p. 292.

¹⁹ Neander, “Functions as Selected Effects: The Conceptual Analyst’s Defense”, p. 180.

evaluative. To take one well-known example: for the teens it is biologically normal to be fertile, but in many situations, it is not good as well.²⁰ Also, the norm should not be identified with statistical normality: although the standards often correspond with statistical averages, they are not the same²¹. Why is it so? I'll make a short digression to show what the problem is. It is easy to imagine a population in which most members have some (the same) dysfunctional trait. For instance, in a population which is exposed to radiation most individuals may suffer from some kind of disease because of harmful genetic mutations. Therefore, it seems that there is no necessary connection between the current functionality of some trait and its frequency in the population. The difference between these two criteria is even more obvious if we analyze examples taken from social context. For instance, we can imagine that most members of a society do not respect the legal rule according to which it is forbidden to consume alcohol while driving. In this case, the person who never drinks would be regarded as deviant. Although it is possible to construct many similar examples, another question is more relevant to the scientist: why, as a rule, the set of traits which most members of a population have coincides with the set of traits which we can call 'functional'? As we have seen, the proponents of proper function say that Darwinian principle of natural selection gives a satisfying explanation of this coincidence.

2. The physiological 'medical model' and the normative capacity of organisms

Boorse's conception of health, as he said many times, is „as value-free as statements of biological function“.²² Three ideas are important for his concepts of health and disease: function, species biological design and normality. In one of his early articles on medicine Boorse determines diseases as “interferences with natural functions”.²³ The question is, of course, what 'natural function' means. At the very core of Boorse's theoretical concept of health is (the concept of) physiological function, which is defined by species design. Physiological medicine is usually understood as a paradigm health discipline and according to Boorse, health vocabulary is physiological one.²⁴ This means that assumptions of the

²⁰ Christopher Boorse, “On the Distinction between Disease and Illness”, *Philosophy & Public Affairs*, 1975, vol. 5, n° 1, p. 52; Neander, “Misrepresenting and Malfunctioning”, p. 111.

²¹ Millikan, *Language, Thought, and Other Biological Categories*, p. 4.

²² Christopher Boorse, “Health as a Theoretical Concept”, *Philosophy of Science*, 1977, vol. 44, n° 4, p. 542.

²³ Boorse, “What a Theory of Mental Health should be”, p. 63.

²⁴ *Ibid.* 62.

physiological ‘medical model’ determine all our thinking and speaking about health and disease. In a sense we can see Boorse’s notion of disease as similar to Bernard’s physiological concept. However, Boorse obviously does not determine the function in terms of pure quantities: according to him, ‘abnormal’ is not “too much or too little” as it is for Bernard.²⁵

Boorse rejects the interpretation of normal as the statistical mean for two reasons: first, there are deviations from the statistical average, which cannot be understood as diseases. Also, some unhealthy states are almost universal, but nevertheless they are still unhealthy.²⁶ Instead, Boorse offers a different interpretation of the statistical criteria for disease: „disease is only statistically species-subnormal biological part-function“.²⁷ The reference class is a class of organisms which share certain functional design. We can see that two criteria are used to define health: the statistical one (normality) and the biological one (the function). Undeniably, both criteria are necessary to understand the healthy state. The link between frequency and adaptive value of one trait is also explained by the theory of evolution: the allele that increases the chances of an organism to survive and to reproduce will become fixed. However, Boorse rejects the notion of proper function and tries to reduce proper function to physiological (CR) function. What he actually does, is to use two inconsistent conceptions of function to define health: he is forced to use the CR function to define the concept of health as a functional contribution to the physiological goals of the organism, and at the same time the evolutionary conception of function to introduce the element of species design (or the biological idea of natural design) in the definition of health. As we have seen, there is no doubt that physiological function is useful in some contexts, but this conception of function is not suitable to serve as a foundation for the concepts of health and disease, precisely because of the lack of the element of normativity. On the other side, the proper function is normative and cannot be assimilated into the first conception.

Boorse determines his view as a population view.²⁸ To understand what this is about I will try to briefly address some important ideas of the so-called population thinking. This type of conceptualization of living world is linked to Darwin’s theory of evolution, although the name comes from later theories. Darwin considered that the species as a

²⁵ Canguilhem, *The Normal and the Pathological*, p. 56.

²⁶ Boorse, “On the Distinction between Disease and Illness”, p. 50.

²⁷ Christopher Boorse, „A Rebuttal on Health“, in: Humber, J.M., Almeder R.F. (eds.), *What Is Disease*, New York, Springer Science+Business Media, 1997, p. 4.

²⁸ *Ibid.*, p. 27.

taxonomic category does not constitute a satisfactory theoretical and empirical framework in which the connection between individual variability and evolutionary changes could be explained, because the species usually does not coincide with the local population. Pre-Darwinian biology was dominated by the typological concept of species, according to which species is a group of organisms which share certain characteristics, in other words, they are instantiations of a constant and unchanging type. So, the great French biologist Buffon says in his *Natural History* (1753): „There is, in nature, a general prototype in each species upon which each individual is modeled... The first animal, the first horse, for example, has been the external model and the internal mold upon which all horses which have ever been born, all those which now exist, and all which will arise, have been formed.“²⁹ In contrast to the typological thinking, which is primarily focused on the constancy of the prototype, the population thinking considers the differences between individuals, i.e. individual variability as crucial for the understanding of the living world. How are these two approaches related to the problem of abnormality? In pre-Darwinian biology, variations among organisms were interpreted as deviations which nature must remove in order to bring a species in harmony with its internal norm. According to Darwin, individual variations have a completely different sense thanks to the principle of competition. Namely, „if all causal transactions are reducible to forces exerted by and applied to individuals competing in a closed system, no high-level, conceptual reasons remain for continuing to assume that variation must be confined within a species-specific limit, or that a species necessarily snaps back to some preestablished norm, or that as one departs from type (whatever that may now mean) organisms are necessarily less fit“.³⁰ Hence we see that the traditional understanding of adaptedness as effect of final cause is replaced by Darwinian relative adaptedness produced by natural selection. What I would like to do now is to consider the question whether this understanding of variability allows for a satisfactory explanation of species design.

Let us see at last what is meant by ‘species design’ and by ‘reference class’. This is very important, because Boorse’s concept of species design refers to typical physiological functioning of organisms, which can be understood as health, while the deviation from this statistically normal functioning is disease. How do scientists use the term ‘design’? ‘Design’

²⁹ Upon Robert J. O’Hara, “Population thinking and tree thinking in systematics”, *Zoologica Scripta*, 1997, vol. 26, n° 4, p. 324.

³⁰ David J. Depew, Bruce H. Weber, *Darwinism Evolving. System Dynamics and the Genealogy of Natural Selection*, Cambridge, Massachusetts, MIT Press, Bradford Book, 1996, p. 132.

mostly denotes “any nonaccidental process by which purposiveness is produced”³¹ and it is usually used to explain the function. However, there is no agreement in biology about what is meant by ‘design’. Nonetheless, it is clear that design is associated with function, moreover with purpose. How does Boorse determine species design and is this notion consistent with evolutionary variability? Is species design a kind of idealization? Boorse claims that statistical idealizations, which represent typical functioning of a trait in the population as reference class are unavoidable in any judgment on health.³² Whether we like it or not, we are faced again with statistical criteria in health analysis. Two important objections can be addressed to Boorse’s approach. First, if this model is valid, we should find only stabilizing selection in nature (or normalizing, as Boorse call it), in which “intermediate individuals do better than the extremes”³³ and individual variations in a population should be decreased. But this is not the case. In the end, it can be expected that people will also be sick in circumstances that are not standard. Second, there is a question to which extent idealizations can help us judge whether someone is sick, since in this statistical context the organism is not seen as a whole with its specificity, but as a collection of characteristics that will be measured relatively to an abstract range of values. The relation between the concrete organism and the population as a reference class is anything but clear, because physiological functions can be performed in many nonstandard ways that we cannot interpret as manifestations of disease.

I would like to summarize my analysis of the connection between function, normality and health. We know that teleology implied normativity: thanks to the notion of purpose we were able to distinguish a healthy organism from the sick one, the normal from the abnormal, just as we were able to distinguish well-built house from the one which is not. Generally, translation from teleological language to the language that does not use teleological terms (and this corresponds to Boorse’s position) is connected with a lot of problems, and this for several reasons. Namely, the explanation of some trait or of the presence of some organ which is expressed in “nonteleological” terms can no longer exhibit causal relationship between the origin of the trait and its presence in the organism except by reference to the so called normal circumstances (or Boorse’s “normal environment”³⁴): for example, chlorophyll is present in green plants because it makes photosynthesis

³¹ McLaughlin, *What Functions Explain*, p. 150.

³² Christopher Boorse, “Health as a Theoretical Concept”, p. 557.

³³ John A. Endler, *Natural Selection in the Wild*, Princeton, New Jersey, Princeton University Press, 1986, p. 17.

³⁴ Boorse, „A Rebuttal on Health“, p. 83.

possible in normal circumstances. Thanks to the reference to normal circumstances we can avoid two sets of difficulties. The first set arises from the relation between the structure and the function performed by this structure (one and the same function can be performed by many different structures). The second set of difficulties is connected to the problem of distinguishing the function from the mere effect. Namely, ‘normal’ can have the meaning of ‘something that happens in most cases’ but also of ‘something that is adaptive’. However, this maneuver does not achieve anything significant, the problem only moves to another plan. On one side, the very fact that we usually associate normality with the normal distribution of some characteristic speaks against the attempt to define a function using normality. On the other side, if ‘adaptive’ should be the second meaning of ‘normal’, then, instead of explanations, we have the uninformative claim that the characteristics of organisms (perhaps) are adaptations.

Because of all this some authors claim that normality is independent from the function and prior to it. In my opinion, Canguilhem’s understanding of health corresponds to this approach. As said before, Canguilhem’s holistic view on health is constructed in opposition to Bernard’s physiological conception of the normal and the pathological “according to which the pathological phenomena found in living organisms are nothing more than quantitative variations, greater or lesser according to corresponding physiological phenomena.”³⁵ Unlike Boorse’s, Canguilhem’s understanding of health is not based on the idea of normality as function, but quite explicitly on the idea of a norm. We have seen that the biggest problem for the naturalistic conception of health is precisely the source of normativity, i.e. the question on the basis of which criteria we should judge the state of an organism. One of the most important ideas in a naturalistic conception like Boorse’s is that health is a theoretical concept that can be defined with the help of fundamental biological theories which determine the goals of the organism. Hence the judgment about a state of organism has nothing to do with the therapeutic judgment. Canguilhem’s standpoint is wholly different. He does not see medicine as a science but as a technique, that is based on the capacities which are inherent to the organism: medicine is “an activity rooted in the living being’s spontaneous effort to dominate the environment and organize it according to its values as a living being. That is “why medicine, without being a science itself, uses the results of all sciences in the service of the norms of life”³⁶.

³⁵ Canguilhem, *The Normal and the Pathological*, p. 42.

³⁶ *Ibid.*, p. 228.

Therefore, the medicine is not a form of knowledge, but a technique that allows us to establish vital norms. The source of the norm is the life itself, more precisely, the interaction between organism and environment. It is therefore clear that the pathological states cannot be analogous to the normal ones and that health is not only absence of disease.

References

- Boorse, C. (1975) "On the Distinction between Disease and Illness", *Philosophy & Public Affairs*, vol. 5, n° 1, pp. 49-68.
- (1976) „What a Theory of Mental Health should be“, *Journal for the Theory of Social Behaviour*, vol. 6, Issue 1, pp. 61-84.
- (1977) "Health as a Theoretical Concept", *Philosophy of Science*, vol. 44, n° 4, pp. 542-573.
- (1997) „A Rebuttal on Health“, in: J.M. Humber, R.F. Almeder (eds.), *What Is Disease*, pp. 1-134. New York, Springer Science+Business Media.
- Canguilhem, G. (1991) *The Normal and the Pathological*, New York, Zone Books
- (1994) *A Vital Rationalist*, New York, Zone Books.
- Christensen, W. (2012) „Natural Sources of Normativity“, *Studies in History and Philosophy of Biological and Biomedical Sciences*, vol. 43, pp. 104-112.
- Cummins, R. (1975) „Functional Analysis“, *The Journal of Philosophy*, vol. 72, n° 20, pp. 741-765.
- Depew, D.J., Weber, B.H. (1996) *Darwinism Evolving. System Dynamics and the Genealogy of Natural Selection*, Cambridge Mass., MIT Press, Bradford Book.
- Endler, J.A. (1986) *Natural Selection in the Wild*, Princeton New Jersey, Princeton University Press.
- McLaughlin, P. (2003) *What Functions Explain. Functional Explanation and Self-Reproducing Systems*, Cambridge, Cambridge University Press.
- Millikan, R.G. (1989) „In Defense of Proper Functions“, *Philosophy of Science*, vol. 56, n° 2, pp. 288-302.

- (2001) *Language, Thought, and Other Biological Categories*, Cambridge Mass., MIT Press, Bradford Book.
- (2002) „Biofunctions: Two Paradigms“, in: A. Ariew ed., *Functions*, pp. 113-143. Oxford University Press.
- Neander, K. (1991) “Functions as Selected Effects: The Conceptual Analyst’s Defense”, *Philosophy of Science*, vol. 58, n° 2, pp. 168-184.
- (1995) “Misrepresenting and Malfunctioning”, *Philosophical Studies*, vol. 79, pp. 109-141.
- O'Hara, R.J. (1997) „Population thinking and tree thinking in systematics“, *Zoologica Scripta*, vol. 26, n° 4, pp. 323-32.
- Okasha, S. (2006) *Evolution and the Levels of Selection*, Oxford, Clarendon Press.
- Wright, L. (1973) „Functions“, *The Philosophical Review*, vol. 82, n° 2, pp. 139-168.

Individuation physique et individuation biologique :
pour un nouveau concept de Nature

1. Une belle antinomie, comme point de départ.

Au paragraphe 30 des *Idées directrices pour une phénoménologie*, Husserl nous présente la thèse de l'attitude naturelle : “ le monde est toujours là comme une réalité ”. Il propose alors la méthode inédite de la phénoménologie idéaliste au paragraphe 32 : s'interdire tout jugement portant “ sur l'existence spatio-temporelle ”. Il faut donc également mettre “ hors-circuit ” toutes les sciences qui se rapportent à ce monde naturel. C'est ainsi que l'on commence à pratiquer ce qu'il nomme “ l'époché phénoménologique ”. Il s'agit de “ mettre en évidence ” que “ la conscience est un être propre (Eigensein) qui, dans son absolue spécificité eidétique, n'est pas affecté par l'exclusion phénoménologique ” (§33). Elle seule nous donne accès à la “ conscience pure ” qui est toujours en même temps “ conscience de ” (§86). La réduction phénoménologique doit être pensée et approfondie dans le sens d'une réduction transcendantale. Il ne s'agit pas simplement de mettre entre parenthèses l'existence spatiotemporelle, mais aussi toute forme d'expérience individuelle et psychologique des états de conscience qui me personnalisent et qui font que ce que je ressens diffère de ce que ressent autrui. Ramenée à son noyau pur, la conscience est co-présente (*mittgegenwärtig*) à toute forme de prédication que l'on pourrait lui attribuer. Il s'agit alors d'isoler à la fois ces vécus, et de réduire toute forme d'objet à être un corrélat intentionnel de ces vécus : “ il n'est point d'objet existant en soi que la conscience et le moi n'atteignent ” (§ 47). C'est ainsi seulement que la conscience peut être “ considérée dans sa pureté comme un système d'être fermé sur soi (*für sich geschlossenerzusammenhang* ” (§49). Comme telle, la conscience pure, la subjectivité idéale, n'a pas de dehors spatio-temporel. Elle est constitutive de tout corrélat intentionnel qui lui est rapportée, mais elle n'est constituée par rien d'autre qu'elle-même.

C'est le premier postulat. Le second réside dans l'idée que nous avons une pure vue immédiate de cette totalité subjective, une vue désincarnée, insensible, une vue aveugle. Il est donc clair, comme le fait remarquer Barbaras,¹ que si le sujet "co-conditionne" l'apparaître, s'il est condition d'apparition du monde, il n'est certainement pas inscrit en lui, et il ne peut pas l'être :

" Le principe est essentiellement différent quand on passe au domaine des vécus en tant qu'essences absolues. Ce domaine est solidement clos sur lui-même, sans pour autant avoir de frontières qui puissent le séparer d'autres régions. Car ce qui le délimiterait devrait encore partager avec lui une communauté propre. Or il est le tout de l'être absolu au sens déterminé que nos analyses ont fait apparaître. Il est par essence indépendant de tout être appartenant au monde, à la nature, et ne le requiert même pas pour son existence (*Existenz*). " (§51).

Entendue de cette manière, la phénoménologie idéaliste exclut tout concept de nature, comme Husserl le conclut très bien lui-même dans cette phrase. Opposons-lui pourtant sans plus de détour une autre thèse, apparemment tout aussi manifeste, mais appartenant à un univers de pensée radicalement contraire :

" Physicalism, the thesis that the physical facts fix all the facts enjoys universal acceptance in the philosophy of biology ".²

Qui refuserait en effet que toute pensée soit faite de neurones, ou mieux encore que toute cellule vivante soit faite d'atomes? N'est-ce pas raisonnable de l'accepter sans sourciller? Derrière cette thèse, comme le rappellent ces deux auteurs qui font partie de la vulgate épistémologique anglo-saxonne, il y a deux postulats qui définissent les principes de la thèse de la "détermination physique" : la complétude et la clôture causale du monde physique. Tous les événements du monde physique peuvent être rapportés au monde physique qui est clos sur lui-même. Le monde physique, dans sa structure, dans son intelligibilité, est par conséquent co-présent à toute forme de prédication que l'on peut

¹ Barbaras R (2003) Vie et intentionnalité, Paris, Vrin, p. 124.

² A. Rosenberg, D.M. Kaplan, "How to reconcile Physicalism and Antireductionism about Biology", *Philosophy of Science*, vol. 72, 2005, p. 43-68.

attribuer à l'un de ses éléments. Il en est condition d'apparition. Mais il n'est lui-même par contre conditionné par rien. Certes on peut ajouter d'autres étages biologiques, puis psychologiques sur cette première base physique. Mais le postulat fondamental sera que toutes propriétés psychologiques d'une pensée, ou encore toutes propriétés biologiques d'une cellule, si elles ne peuvent pas être réduites à des propriétés physiques, en dépendent essentiellement. C'est donc du monde physique qu'il faut partir, non pas seulement pour les comprendre, mais pour les expliquer scientifiquement.

Comment éviter de tourner en rond dans cette antinomie ? *Selon la thèse*, la conscience est co-prédicat de toute prédication sans pouvoir être prédiquée par rien ; *selon l'antithèse* c'est le monde physique qui acquiert ce statut, et la conscience n'est qu'un épiphénomène survenant/émergent, même s'il n'est pas directement dérivé des lois de la physique.

2. Trois thèses bergsoniennes pour sortir de l'antinomie.

* La première thèse célèbre, qui a été commentée par de multiples auteurs, dont Deleuze et Barbaras, concerne la relation entre la perception et la matière. Nous allons partir d'elle : la nature est “ un ensemble d'images ”, mais ces images existent en soi, la perception est elle-même une image, mais c'est une image “ privilégiée ” dont toutes les autres images “ réfléchissent l'action possible ”.³ Voilà pourquoi en un sens, le cerveau est condition de l'image totale, alors qu'en un autre sens, il en est une partie. Comment mieux comprendre ce paradoxe ?

Par le premier élément de la thèse, Bergson rejette la division kantienne entre phénomènes et chose en soi. Ce qu'il défend n'est en rien un perspectivisme. Il s'en explique dans *l'avant propos*. L'objet est bien une image, mais c'est “ une image qui existe en soi ”. Par le second élément, il s'en prend au subjectivisme : c'est en elles que nous percevons les choses, elles réfléchissent notre action possible sur elles. Ce n'est donc pas en nous. Les choses sont comme une sorte de miroir qui reflète l'action que nous pouvons

³ Henry Bergson, *Matière et mémoire*, Paris, Alcan, 1896, p. 21.

avoir sur elles. La perspective que nous allons engager sort simplement d'un commentaire de cette citation de Bergson :

“ Aucune doctrine philosophique ne conteste que les mêmes images puissent entrer à la fois dans deux systèmes distincts, l'un qui appartient à la science, et où chaque image, n'étant rapportée qu'à elle-même, garde une valeur absolue, l'autre qui est le monde de la conscience, et où toutes les images se règlent sur une image centrale, notre corps, dont elles suivent les variations ”.⁴

** La seconde thèse, qui ne peut se comprendre que par cette citation, est que la science pense. Elle ressaisit les choses telles qu'elles sont, et non pas simplement telles qu'elles nous apparaissent.

Autrement dit, la science, ou plus modestement une théorie scientifique n'est pas qu'un ensemble de représentations. Il y a un engagement ontologique dans n'importe quelle forme de théorie scientifique.

Cette thèse est évidemment contraire à celle défendue par l'idéalisme husserlien, mais aussi par l'ontologie phénoménologique heideggerienne. Elle n'a évidemment rien de commun non plus avec le constructivisme idéaliste de Cassirer.

*** La troisième thèse, lointaine héritière d'Héraclite, d'Anaximandre, ou encore de Lao Tseu, c'est qu'il n'y a pas de priorité de l'être sur le devenir. Mais il faut la comprendre dans la rigueur de pensée qu'elle exige. Nous ne pouvons pas être quittes de cette thèse en proposant une ontologie du devenir. Ce serait en effet, encore un coup affirmer qu'il y a un être du devenir. Pour éviter de tomber dans ce piège, il faut se méfier des idées négatives.

Il faut se rendre compte que le principe du tiers exclu qui gouverne la logique traditionnelle, nous oblige à penser l'être sur bord de non être, et l'ordre sur bord de désordre. Nous nous proposons de le lever. Il faut que ce qui n'est pas, soit pourtant en même temps quelque chose. Ce qui n'est pas n'est pas du rien, c'est un autre plein. Et ce qu'il faut c'est une métaphysique qui nous fasse comprendre comme le non être est un autre être, le désordre un autre ordre. Une métaphysique qui mette au premier plan ce bougé des choses, ce procès, cette altérité.

⁴ *Ibid.*

C'est bien la métaphysique que Bergson appelle de ces vœux selon nous, quand il propose dans *L'évolution créatrice* de comprendre l'ordre géométrique comme une substitution de l'ordre vital et non comme sa suppression. Il reprend ainsi l'intuition de *Matière et mémoire* : la matière n'est pas du rien. Elle est réelle, comme la vie, comme l'esprit peuvent l'être. Et le mot " réel " est plus important que le mot " être ".

Regroupons donc à présent ces trois thèses et revenons à la citation de Bergson. Nous allons en fournir une interprétation, qui est en même temps une reconstruction de la position bergsonienne. Pour introduire cette reconstruction, nous allons nous-mêmes nous servir d'une image.

Revenons à la figure de la conscience. La conscience est un photographe qui transforme la réalité en clichés. Il ne nous viendrait pas à l'idée de dire que la réalité n'existe pas, qu'elle n'est pas toujours déjà là, puisque le photographe en prend des clichés. Pourtant sur ces clichés, rien de la réalité n'apparaît directement, puisque tout ce que nous savons du monde extérieur, nous ne le savons ainsi que par nos sens. Nous ne savons des choses, que ce que nous percevons d'elles. Il est donc tentant d'en conclure que leur être, ce qu'elles sont en tant que telles, nous échappe.

Si cette position était juste, prolonger la perception par la science ne changerait rien au problème. Il faudrait juste ajouter que nous n'avons des informations sur les choses, qu'à partir de ce que nous connaissons d'elles. Comme le disait Kant, nous ne pouvons connaître que des phénomènes et non la chose en soi. Selon lui, notre connaissance n'est pas nouménale. Nous ne disposons pas d'un entendement divin. Donc la réalité échapperait à la science, comme elle échappe à la perception et il serait absurde de parler d'un engagement ontologique des théories scientifiques.

Et cependant nous voyons bien qu'en suivant cette voie, nous sortons de la position bergsonienne, pour qui la science pense et pour qui la réalité n'échappe pas au savant qui fait ainsi une expérience à la fois physique ou mathématique, mais en même temps métaphysique. Comment réveiller l'intuition bergsonienne ? Il faut pour cela imaginer **une expérience de décorrélation**. A la fois c'est une décorrélation, mais en même temps nous sommes en mesure d'en faire l'expérience. Revenons à notre image du photographe, pour nous faire bien comprendre. **Redoublons-le simplement**. Plaçons un second photographe à

côté du premier. Il est à son monde, comme le premier. Mais sur son cliché, il voit apparaître la différence entre les photos prises par le premier et le modèle.

Gageons que c'est ce que fait l'expérience scientifique. Ce n'est pas une simple perspective, comme l'affirmaient Nietzsche, ou dans un autre genre Cassirer. Elle engage un redoublement d'immanence. En un sens, d'un certain point de vue, nous sommes à nous-mêmes, lorsque nous percevons, imaginons ou encore connaissons. Nous sommes dans le plan de conscience qui n'est en rien réductible à un phénomène physique. Nous sommes dans le monde de nos pensées, dans un monde subjectivé. Mais en un autre sens, la science est comme le second photographe. Elle fournit une image de la réalité, non pas telle que nous nous la représentons, mais telle qu'elle se rapporte à elle-même. ***La réalité, telle qu'elle se rapporte à elle-même, dans sa pure immanence, est donc susceptible d'être simulée par la pensée. Ce n'est qu'une image qui existe en soi, ce n'est pas une chose en soi.*** A la fois monde et image, à la fois absolue et événement. Mais puisque cette image existe en soi, à travers le redoublement ainsi effectué dans la pensée, c'est bien aussi le redoublement réel entre le plan d'immanence et le plan de conscience qui est exprimé.

Enfin, dernier point, mais pas le moindre. Une expérience de décorrélation n'est pas obtenue par un cadrage, par une simple réduction, ***mais par un redoublement, par une extension du cadrage.*** Il faut pour cela que la science ne retranche pas simplement quelque chose à la conscience perceptive, mais qu'au contraire elle ajoute, elle redouble, elle étende ou encore élargisse le regard. ***Elle ajoute le monde physique et formel au monde de la conscience.*** C'est cette ouverture ontologique qu'elle pratique que le philosophe doit réfléchir, et c'est pour désigner ce redoublement d'immanence que nous avons choisi le beau nom de “ ***Nature*** ”. Le concept de Nature n'est pas scientifique. C'est une catégorie philosophique. La science engage ontologiquement. Elle montre quelque chose que le philosophe a pour tâche de réfléchir. C'est donc une nouvelle extension que pratiquera ce troisième photographe pour avoir accès au concept de Nature. Ce n'est pas une simple expérience de décorrélation, comme celle du savant. Il ne lui suffit pas de placer la bulle de la conscience dans celle du monde physique. Il va s'interroger ***sur le mouvement de bascule*** qui nous fait continuellement passer de l'un à l'autre. Cette interrogation requiert une philosophie du procès et du devenir. C'est elle que le troisième photographe va tenter de ressaisir.

3. L'engagement ontologique d'une théorie scientifique.

Quel est le truc qui permet de pratiquer cette expérience de décorrélation ? Ce truc a déjà été aperçu pour des raisons différentes par Barbaras et Deleuze. Nous allons supposer que ce redoublement apparaît déjà dans la pensée elle-même. Nous insistons simplement sur le rôle de la science qui n'est pas une pensée comme les autres.

* Il faut comprendre la pensée à la fois comme co-prédicat de toutes ses prédications, et pourtant en même temps partir du principe que ***ce qu'est*** la pensée n'est pas ce qu'elle ***fait***. Il faut partir du principe que la pensée est “ non coïncidence à soi ”, comme la grenouille de la fable de Lafontaine. Mais cette propriété de ne pas coïncider à soi ***qui rend la pensée vivante et même matérielle*** est en même temps ce qui place d'emblée celle-ci dans un monde qui n'est pas le sien. La distance entre ce qu'elle est et ce qu'elle fait, fait aussi de la pensée un événement dans un autre monde dont elle est, mais qui n'est pas le sien. Elle est incluse en lui, elle procède de lui sans pourtant perdre son statut subjectif.

Dans toute forme de pensée, on peut repérer cette propriété qui fait d'elle en même temps un ***acte*** en plus d'une représentation, d'une description, d'un sentiment, d'une connaissance, d'une prescription, etc. C'est ainsi notamment que nous réinterprétons le *cogito* de Descartes : chaque fois que nous pensons, nous savons que nous existons, ***c'est donc dans l'acte de pensée que nous le savons***. C'est cet acte qui est premier par rapport à toute forme d'intuition ou encore d'évidence.

** Mais nous allons avancer une thèse supplémentaire. Si toute forme de pensée est ainsi relationnelle, si elle engage, en tant qu'action psychique l'existence d'un monde qui n'est pas le sien, la pensée scientifique est une pensée systémique. ***Seule la pensée scientifique a le pouvoir de redoubler les sensations en un monde physico-formel dans lequel nous sommes mais qui n'est pas le nôtre***. Elle seule a le pouvoir de faire apparaître dans la conscience, quelque chose qui se réfère à soi, et non pas seulement à nous qui le pensons. La question que nous allons donc d'abord nous poser ici est : comment s'y prend elle ? Et nous allons nous focaliser sur la physique pour répondre.

*** Ensuite, à travers le redoublement qu'elle opère, nous avons dit que la pensée scientifique montre ce que le philosophe doit réfléchir. Elle montre un monde physico-

formel dédoublé du nôtre. C'est le processus d'objectivation que nous allons décrire. Pourtant nous allons voir que ce qui se produit dans la pensée se produit également dans l'objet que la pensée analyse. Cet objet objectivé lui-même se dédouble : le monde physique se dédouble en mondes biologique et psychique. C'est cette objectivation qui garantit que nous sortons de la simple sphère des vécus de conscience. Ce que nous avons pensé en analysant la relation entre les plans d'immanence et de conscience, la science le fait, la science le montre dans le plan d'immanence lui-même.

Dans le dédoublement entre le monde physico-formel et le monde biologique se reflète et s'exprime le redoublement initial entre les plans de conscience et d'immanence, objectivant ainsi ce qui n'était que subjectivement contenu initialement dans l'univers de simulation de la pensée humaine. C'est ainsi que nous pouvons avoir la garantie que ce que nous avons aperçu dans la pensée se trouve également dans les choses : une ontologie du dédoublement, de la finitude du réel qui est toujours à distance de lui-même, une ontologie qui se mue en philosophie du devenir.

4. De la complexité physique au dédoublement du monde physique en monde biologique.

* Cette manière de lancer le problème donne immédiatement un angle nouveau pour traiter la question de la complexité physique. Nous partons en effet du principe que les systèmes physiques complexes ne sont justement pas des ensembles que l'on peut définir et réduire à la somme des éléments qui les composent. Ils sont au contraire définis par des structures circulaires que rien dans l'axiomatique des théories physiques traditionnelles ne permet d'intégrer. Nous allons nous expliquer plus en détail sur cette question. Il est clair cette fois que notre source d'inspiration n'est plus Bergson, mais Simondon. C'est de *L'individuation et sa genèse physico-biologique* que vient le point de vue que nous défendons. Nous allons seulement rendre plus explicite et précis un nombre de formulations qui s'y trouvent déjà. Pour ce faire, nos deux autres sources d'inspiration profonde, sont d'abord les travaux d'Ilya Prigogine, Isabelle Stengers, et Grégoire Nicolis (*La nouvelle Alliance, Le temps retrouvé, Exploring Complexity*), et ensuite à des titres

divers ceux de Stuart Kauffman (*Investigations*), Francisco Varela (L'autonomie du vivant), Henri Atlan (*Le cristal et la fumée ; La fin du tout génétique*), Francis Bailly et Giuseppe Longo (*La singularité physique du vivant*), Denis Noble (*La musique de la vie*), Ana Soto et Carlos Sonnenschein (*The Society of Cells*).

** Nous allons d'abord défendre une première thèse : dans tout système physique si ample, nous voyons déjà apparaître le redoublement entre le monde de la pensée et le monde physique, justement dans la mesure où ce système est *théorique*. Nous allons préciser ici le sens hautement non idéaliste que ce mot a pour nous.

Il est important en effet de bien comprendre d'abord qu'une théorie scientifique n'a rien d'empirique. Elle ne consiste pas en un ensemble d'énoncés qui entre en adéquation avec un ensemble de faits. Il faut tordre le cou à cette conception de la vérité simpliste issue du Moyen-Age et réactualisée par Russell et Tarski. Tout au contraire, chaque physicien sait bien que c'est contre la conscience empirique qu'une théorie scientifique se développe et non pas simplement dans son prolongement. Comme l'écrivait Bachelard, l'objet physique est un *surobjet*. Comme le disait Popper, une théorie scientifique n'est pas vérifiée par l'expérience. Comment pourrait-elle l'être puisqu'elle contient des énoncés nomologiques ? Une théorie est scientifique si elle ouvre une liste de *falsificateurs virtuels*, si elle engage des procédures d'expérimentations, qui vont conduire peu à peu à transformer et à étendre cette théorie. Il n'y a donc jamais aucun fait empirique à quoi une théorie se rapporte. Un atome n'est pas un soleil. Les électrons ne sont pas des planètes. Une particule n'obéit pas à la loi du *hic et nunc*. Un système dynamique peut être complètement déterminé et pourtant en même temps complètement imprévisible. Comment donc s'y prenait Pasteur pour montrer l'agent étranger responsable d'une maladie, sinon avec l'expérience du vase en col de cygne ? Ce qu'il montrait dans le vase n'avait rien à voir avec un fait empirique, c'était un fait expérimental, entièrement relié au dispositif technique imaginé et conçu par le savant. Ce processus commence dès le départ avec l'expérience du bateau des Dialogues de Galilée. C'est le sens commun aristotélicien qui dit que lorsque le bateau avance à vitesse constante, la pierre jetée par le mousse en haut du mât va tomber dans l'eau, et c'est le génie théorique et expérimental de Galilée qui lui répond qu'elle tombe au contraire au pied du mât. *Nous voyons bien ainsi qu'une théorie scientifique n'est pas seulement un ensemble de représentations mais plutôt un*

geste qui nous fait constamment basculer au-delà du monde de nos représentations empiriques. Par une théorie physique nous ne sommes pas simplement dans notre monde, nous sommes toujours en même temps d'un monde qui n'est pas le notre et qui nous oblige en permanence à élargir notre regard.

Mais ensuite dire que la physique développe une approche théorique dans le regard qu'elle jette sur les systèmes physiques, c'est précisément affirmer qu'elle propose une expérience de décorrélation, une expérience d'objectivation. Dans et à travers cette expérience, ces systèmes n'apparaissent plus comme les corrélats intentionnels d'un vécu de conscience. Cette expérience est au contraire une expérience de rupture de corrélation simulée dans la pensée. Dans la pensée, nous analysons des systèmes qui se réfèrent à eux-mêmes et non simplement à nous qui les observons et les interrogeons. C'est à présent les vécus de conscience qui sont comme les corrélats d'une intentionnalité seconde qui n'est plus la nôtre, mais celle ***du point de vue du système*** qui se dédouble du ***point de vue de l'observateur***. Telle est à notre avis l'origine de l'hypothèse du Démon de Laplace. Elle est conçue pour exprimer ce redoublement. Sauf qu'il ne faut pas oublier qu'il s'agit d'une simulation et que c'est toujours du point de vue de la pensée humaine, au sein d'une expérience humaine, que cette distinction de points de vue apparaît. Dès que nous l'oublions, nous tombons dans la tentation d'un rationalisme essentialiste comme celui de Leibniz qui oublie que nous ne sommes inclus dans un monde qui n'est pas le notre que parce qu'en même temps nous avons des images de ce monde en nous, de telle sorte qu'il est lui aussi inclus en nous, comme nous sommes inclus en lui et que ce qui est premier n'est ni nous, ni lui, mais le mouvement à travers lequel nous basculons sans cesse de l'un à l'autre.

Sauf que dans les systèmes physiques classiques cette expérience de décorrélation est idéalisée. Dans les systèmes physiques classiques en effets, ***les états physiques peuvent être rapportés entièrement à leur structure de détermination***. Si nous prenons l'exemple simple de la seconde loi de Kepler en mécanique céleste, elle dit que les surfaces des triangles balayées par une planète qui tourne autour du soleil en un temps donné sont toujours les mêmes, alors que tous les états physiques du système changent. La planète change de position, elle change de vitesse, la force d'attraction elle-même change de valeur,

car la trajectoire est elliptique. Pourtant on peut rapporter tous ces états physiques, tous ces changements à une structure invariante qui prend la forme d'un principe physique : la conservation de l'énergie mécanique. L'opérateur H qui permet de rapporter ces changements à la structure invariante est l'hamiltonien du système :

- $dp/dt = - \partial H / \partial q$
- $dq/dt = \partial H / \partial p$

C'est ainsi que le système physique commence à se référer à lui-même à travers le regard que nous portons sur lui. Grâce à l'opérateur hamiltonien, ***il est clos sur lui-même***. Il est défini entièrement par la forme qu'il prend dans l'espace de ses états. C'est un système fermé. Disons-le autrement : il n'existe aucun de ses états qui ne puisse être entièrement rapporté à sa structure. Ce système est défini par cette relation interne, immanente, entre les états et la structure qui fait qu'il se réfère à lui-même et qui fait que nous pouvons penser aussi que la structure est comme un principe de Raison dont les états ne sont que les manifestations qui l'expriment. Un système fermé de ce genre est donc aussi ***conservatif***. Il est atemporel. Il est défini par ses principes d'invariance, grâce à l'opérateur :

- $H = Cte; dH/dt = 0$

Dans l'exemple que nous venons de prendre, l'opérateur est rapporté à l'énergie mécanique, mais il peut évidemment être rapporté à d'autres formes d'invariants physiques. Maintenant nous allons supposer que les systèmes conservatifs de ce genre sont des systèmes idéaux. Cela ne signifie pas que tout est faux dans la mécanique classique. Tout est simplement idéalisé. Et il ne faut pas l'oublier. Nul autre que Simondon n'exprime ce point avec autant de pertinence. Donnons-lui la parole :

“ Le déterminisme resterait applicable comme cas-limite lorsque le système n'est le théâtre d'aucune individuation, c'est-à-dire lorsqu'aucun échange ne s'accomplit entre structure et énergie, venant modifier les structures du système. ”⁵

⁵ George Simondon, *L'individu et sa genèse physico-biologique*, Paris, PUF, 1964, p. 146.

*** Nous allons à présent nous servir d'un concept de Simondon pour indiquer comment des circularités commencent à apparaître au sein des systèmes physiques complexes, **le concept de potentiel structural**⁶ (1964, p. 75) qui a souvent été bien mal compris par ses lecteurs et interprètes.

Par “ potentiel structural ” en effet, il ne faut pas entendre simplement la capacité qu'a un état physique de changer à l'intérieur d'une même structure, mais plutôt la capacité d'un système physique à changer de structure. C'est cette capacité que Simondon propose d'étudier, lorsqu'il décide de s'intéresser à ce qu'il nomme le processus d'individuation. L'individuation physique est un processus. Il a donc un commencement dans le temps. Il existe **des conditions** qui font qu'un potentiel de transformation apparaît à l'intérieur d'un système physique complexe. ***Mais en même temps, on ne peut pas analyser ces conditions simplement depuis ce système physique lui-même.*** Le système est donc ici défini par le fait qu'il est toujours déjà pris dans une relation plus originaire que lui. Cette relation, ou encore cette “ contrainte ” comme on dit aujourd'hui (Prigogine, Nicolis, 1989), est d'abord une **contrainte topologique**. ***Un système physique complexe est d'abord défini par ses contraintes topologiques qui ne sont pas comprises dans l'axiomatique permettant de caractériser sa structure physique.*** Telle est la première thèse qui explique qu'il s'individue avant d'être un individu. Ce que dit alors Barbaras de la vie vaut en réalité déjà pour un tel système physique : il est défini comme “ un être qui a son essence en dehors de lui-même ” (2004, p. 192), un être détotalisé, **un ouvert**. Il n'a pas simplement un dedans qu'un hamiltonien pourrait décrire. Il est ouvert à une relation dedans/dehors. Si nous nommons ce système X, nous avons alors l'usage ici d'un opérateur d'un genre nouveau qui définit X par la relation qui l'intègre en même temps dans une structure plus grande, et qui n'est ainsi qu'une relation entre lui et lui-même. Ce qu'il est n'apparaît pas sur le fond de ce qu'il n'est pas, comme lorsqu'on définit un objet. Ce qu'il n'est pas, fait au contraire partie de ce qu'il est. Il n'est pas défini par ce qu'il est. C'est plutôt la relation entre ce qu'il est et ce qu'il n'est pas qui définit ce qu'il est. Cette relation ne s'explique pas par son être. C'est plutôt de cette relation que vient son être lui-même. Pour le cristal, ces contraintes externes seront des changements de température et de pression. Pour les convections de

⁶ Ibid.

Bénard, ces contraintes seront le nombre de Rayleigh et le nombre de Prantl, mais le premier au moins suppose un gradient de température, un flux thermodynamique exprimé par le fait qu'un liquide est placé entre une source chaude et une source froide.

Ensuite dans un système complexe, *ce potentiel de transformation peut être activé*. L'activation d'un potentiel de transformation, correspond à l'émergence d'une contrainte globale interne, dans le langage de Prigogine. L'opérateur à travers lequel cette activation s'effectue est un opérateur chronologique avant/après. Dans un tel système, le temps-devenir n'est plus annulé. Il agit. La condition idéalisante : $dH/dt = 0$ est donc supprimée. Typiquement c'est le processus de cristallisation annoncé par le fait que les valeurs de volume et d'entropie se mettent à diverger aux points critiques de pression et de température dans une transition de phase du premier ordre. Ou encore c'est l'émergence des rouleaux de convection de Bénard. Cette contrainte externe émergente *est le résultat de l'individuation du système*. Elle n'est pas une disposition, ou encore une qualité cachée, comme pour les métaphysiciens essentialistes. *Elle n'est que le résultat de la relation entre ce qu'est un système physique et ce qu'il opère*. Il n'est plus défini par ce qu'il est. *Il n'est plus objet, mais agent*. Ce qu'il est dépend maintenant de ce qu'il fait et son être commence à avoir valeur d'être. Ce n'est plus son processus qui dépend de sa structure, mais sa structure qui dépend de son processus. Avant, pas de cristallisation, pas de rouleau. Après, de la cristallisation. Des rouleaux. La relation chronologique dans lequel il est pris n'est donc pas dérivée de sa structure de détermination.

Nous affirmons alors dans la lignée de Simondon qu'un système physique complexe, comme peut l'être un système thermodynamique ouvert loin de l'équilibre, est un ensemble d'états physiques qui ne coïncide plus avec sa structure de détermination, puisque ces états dépendent à présent également des contraintes topologiques et chronologiques à travers lesquelles cette structure se transforme. On ne peut donc pas expliquer ces transformations par le système lui-même. *Il faut l'élargir. Il faut ajouter ces contraintes qui rendent ce système circulaire, qui font de lui un agent et non plus un objet*.

**** Posons-nous alors une dernière question pour arriver au bout. Comment comprendre la transition qui fait de nous, non plus un cristal ou une flamme, mais des êtres

vivants ? Simondon avait proposé une réponse. Il affirmait que le régime d'individuation changeait chez un être vivant. *Quelle est donc la condition pour qu'un potentiel vital puisse se former ?* Il faut, disait Simondon, que l'individuation soit mise en scène, mais qu'elle soit en même temps le théâtre à l'intérieur duquel on la met en scène.

Dans un système physique, l'individuation est un état. Mais c'est un état d'exception. C'est un état qui s'explique par des contraintes topologique et chronologique qui ne dérivent pas de l'axiomatique permettant de caractériser la structure du système. Un système complexe envisagé de cette manière *est donc structurellement incomplet*. Il existe dans la théorie qui le définit des formules qui ne peuvent être dérivées de l'axiomatique de cette théorie.

La question est donc de savoir si on peut élargir le regard. Peut-on faire des contraintes topologique et chronologique des formules qui puissent être dérivées d'une axiomatique plus large, faisant ainsi des systèmes physiques complexes comme le cristal ou la flamme un cas particulier qui dérive de leur loi générale, *un cas idéalisé où le potentiel vital n'est pas encore formé*.

La réponse à cette question n'est plus dans Simondon. Il en annonçait pourtant l'esquisse : la condition pour qu'un tel potentiel se forme est "une récurrence de causalité".⁷ Peut-on donner un sens plus explicite à cette formule obscure ? Nous croyons que oui.

Si nous y parvenons, il donnera aussi une épaisseur conceptuelle à l'hypothèse proposée par Bailly et Longo⁸ (2006) selon laquelle, pour comprendre ce qu'est un organisme vivant et comment il évolue, il faut se placer dans un espace critique. Un système physique complexe peut changer d'état. Le point autour duquel il change d'état est un point critique assorti de certains exposants que l'on peut calculer. Mais un système biologique passe continuellement d'un point critique à un autre. Il ne se meut plus dans un espace physique. Il se meut dans un espace critique.

⁷ Simondon, *op. cit.*, p. 160.

⁸ F. Bailly, G. Longo, *Mathématiques et sciences de la nature. La singularité physique du vivant*, Paris, Hermann, 2006.

Nous donnerons aussi une valeur aux hypothèses audacieuses de Soto et Sonnenschein.⁹ Pour eux en effet, ce qui caractérise un organisme n'est pas la reproduction, mais la prolifération avec variations. Ce n'est pas non plus simplement l'agentivité, mais la motilité. Il faudrait revenir plus longuement sur ces deux hypothèses. Mais notons que la prolifération avec variations n'est pas non plus un simple processus chimique de croissance. Notons également que la motilité est la capacité d'être actif et non pas seulement celle d'être agent.

Enfin cette hypothèse est également en relation avec le concept proposé par Varela de "clôture organisationnelle".¹⁰ A travers ce concept, il est question du passage d'une structure physique causale à un univers où cette structure physique causale se dédouble en monde normatif. Un organisme n'est plus qu'un ensemble d'états physiques, mais il est en même temps défini comme un ensemble de contraintes impliquant la relation entre lui et son milieu environnant.

Revenons donc sur notre point de départ. Un système biologique est du monde physique. Il se caractérise par un ensemble d'états physiques rapporté à une structure. Pourtant en même temps il est défini par l'action de ses deux opérateurs dedans/dehors et avant/après couplés. Il s'individue. Comment maintenant l'individuation pourrait-elle se redoubler ? La réponse nous semble simple dans son principe, il faudrait trouver une nouvelle forme de clôture au niveau du couplage entre les opérateurs eux-mêmes. Il faut que la clôture organisationnelle du vivant vienne redoubler sa structure physique.

Mais par ailleurs la structure physique d'un système de ce type est définie par son incomplétude, par le fait que dans sa théorie il existe des formules qui ne peuvent être rapportées à son axiomatique. Le redoublement dont il s'agit par conséquent doit pouvoir se situer à ce niveau : l'hypothèse est tout simplement que le couplage de ces deux opérateurs soit appliqué indéfiniment à lui-même, *qu'il forme une suite de boucles récursives*, dont on va analyser les points fixes de la dynamique.

Il suffit donc que les deux opérateurs d'ouverture topologique et d'agentivité soient à la fois internalisés et appliqués à eux-mêmes un nombre indéfini de fois pour qu'ils

⁹ C. Sonnenschein, A. Soto, *The Society of Cells*, Oxford, Bios Scientific Publishers, 1999.

¹⁰ Voir F. Varela, *Autonomie et connaissance*, Paris, Seuil, 1989.

forment un monde biologique. C'est ainsi qu'ils se transforment en théâtre à l'intérieur duquel les processus biologiques se produisent. En quoi consiste donc ce théâtre ? En un monde qui n'est plus causal, qui n'est plus actuel, ***mais qui devient normatif et virtuel, un monde à la fois vital et vécu.***

Au sein de ce théâtre, les différences dedans/dehors sont internalisées. Elles prennent d'abord la forme d'une mémoire hétérogène. Ce qui rend selon nous la mémoire d'un organisme structurellement hétérogène, c'est le fait qu'il n'ait pas un passé, mais plusieurs. Sa géographie entre en phase avec son histoire. Le principe fondamental qui gouverne cette hétéronomie c'est l'hérédité. L'organisme est défini par la tension entre sa mémoire et la mémoire de l'espèce. C'est même selon nous un instrument de mesure : plus cette tension est grande, et plus un être vivant est organisé. Et cette tension prend une forme concrète et tangible par le fait que ce n'est pas dans les mêmes lieux que ces deux mémoires sont repliées dans l'organisme. Mais dans et entre ces deux termes extrêmes, l'hétérogénéité joue encore. Ainsi par exemple, l'identité physiologique d'un organisme n'est pas acquise pour une colonie, comme elle l'est pour ce que l'on a coutume de nommer un organisme véritablement individué. La ligne de partage entre les deux est visible chez les végétaux, comme l'a montré Francis Hallé. Dans notre perspective, plus un organisme est individué, plus il est dissocié de son hérédité, ce qui contribue à renforcer la ligne de partage entre cellules germinale et somatique ; mais il ne saurait jamais l'être tout-à-fait. Scott Gilbert et Thomas Pradeu viennent de montrer récemment le rôle joué par la symbiose dans l'immunité et le développement des animaux développés.

En sens inverse, l'hérédité d'un organisme n'est jamais simplement son hérédité génétique. Plus l'organisme est individué et plus les facteurs héréditaires liés à des critères développementaux vont prendre de l'importance. La gestion de la relation entre l'organisme individué et son environnement donne une importance cruciale à des modes de transmission héréditaire et à des modes de développement qui ne sont plus entièrement dictés par une hérédité génétique indiquant l'inclusion de l'organisme au sein de l'espèce à laquelle il appartient. Les phénomènes d'empreinte parentale ou encore de paramutations en sont des exemples connus aujourd'hui.

Ensuite premièrement l'action d'un être vivant n'est pas seulement celle d'un agent, comme tel est déjà le cas pour un système physique complexe qui s'individue. " L'agentivité ", la relation avant/après est internalisée. Elle se transforme en " motilité ", dans la mesure où tout organisme vivant se définit *par l'écart, la tension entre l'action effective et l'action simulée*, qui vient de l'écart entre perception et désir. Il n'est donc pas étonnant, comme l'a noté Franck Varenne que la simulation soit aussi importante pour comprendre l'activité biologique. Elle n'est pas un adjuvant, un succédané de théorie. Elle fait ici partie de la théorie à part entière. C'est cet écart entre action réelle et action simulée de par ce processus d'internalisation de l'agentivité, qui transforme l'action en but et son accomplissement en fonction. Ce n'est pas le but qui explique ce processus d'internalisation, c'est plutôt ce repli sur lui-même de la ligne causale qui explique l'émergence d'une action orientée vers des buts. A travers elle l'après est placé avant l'avant. Il devient normatif, il l'oriente.

Deuxièmement, l'hétérogénéité structurale est redoublée en hétérogénéité normative, par couplage dedans/dehors et avant/après. L'individuation biologique ne naît pas simplement par le fait que l'être d'un système biologique a valeur d'être, qu'il est tendu vers un but. *Elle naît de l'internalisation des valeurs contraires* : s'adapter et être adapté ; se développer et vieillir, être malade et être en bonne santé, être vivant et mourir. Ce n'est pas simplement que ce que n'est pas le vivant fait partie de lui, ce n'est pas uniquement qu'il s'organise en relation avec son milieu, comme dans la théorie de l'autopoïèse. C'est plus profondément que la norme à laquelle il obéit n'est pas seulement la sienne. De cette tension vient le concept de " normativité ". La santé, la vitalité, comme l'avait déjà reconnu Canguilhem et sans doute Nietzsche avant lui vient du fait que cette tension soit maintenue. Tel est en effet selon nous le sens profond du concept de " volonté de puissance ".

Et la différence avant/après fait que ce que sont ces valeurs contraires diffère de ce qu'elles font. C'est cette différence entre ce qu'elles sont et ce qu'elles font qui définit véritablement ce qu'elles sont. Elles sont *normatives*. Ce qu'elles ne sont pas n'est pas rien, c'est une autre allure de vie, une allure passive. Ce qu'elles sont n'est pas plein, c'est ce que Soto et Sonnenschein nomme leur motilité, des contraintes propulsives, leur allure active. Dire qu'elles sont normatives, c'est dire que l'être de ces valeurs dépend structurellement de leur agir, voilà pourquoi ce potentiel vital activé produit du nouveau, il

est créateur. Mais le fait qu'il puisse ne pas en produire, le fait de sa passivité, cette altérité à l'altérité fait encore aussi partie de lui-même, est essentiel, structurel à l'itération de la créativité.

Nous ne souhaitons pas prétendre donner les détails scientifiques des processus que nous indiquons. Même pas la théorie. Nous nous contentons d'en dessiner l'ontologie. Si notre hypothèse ontologique est bonne, il y a une rupture de corrélation entre mondes physique et biologique à l'intérieur d'un système biologique. C'est cette rupture qui fait que l'organisme est à son monde tout en étant dans un monde physique. S'il est possible un jour à un savant de la simuler vraiment – et ce jour est proche à notre avis- nous verrons le mouvement de bascule initié par le geste d'objectivation de la théorie scientifique *redoublé dans l'objet même qu'elle analyse*. L'incomplétude du monde n'y apparaîtra donc pas que dans son geste, mais dans l'objet qu'il dégage à travers le geste de simulation qu'il propose. Nous pourrions donc bien en conclure que le monde a du sens, un sens pluriel, habité, inhumain, pour nous autres êtres humains, trop humains. Mais nous concluons aussi que cette inhumain est à notre portée, nous en retrouvons la trace en nous-mêmes. Nous pouvons faire alliance avec lui de telle sorte qu'il puisse enchanter et féconder notre pensée de son inhumanité. Il suffit pour cela comme le fait remarquer Stengers, que nous enfourchions notre balai de sorcière.

Bibliographie

- Adams B., Lipson H. (2009) "A Universal Framework for Analysis of Self-Replication Phenomena", *Entropy* 11, no. 2: 295-325.
- Ameisen, J.C. (2003) *La Sculpture du vivant*, Paris, Seuil, 2003.
- Anderson, P. (1971) "More is different", *Science*, 177, 4047, pp. 393-96.
- Armstrong, D. (2002) "Vérifacteurs pour des vérités modales", *Revue de Métaphysique et de Morale*, vol. 4, n°36, pp. 461-477.
- Armstrong, D. (2004) *Truth and Truth Makers*, Cambridge University Press.
- Atlan, H. (1979) *Entre le cristal et la fumée*, Paris, Seuil, 1979.

- Atlan, H. et alli (1986) "Emergence of classification procedures in automata networks as a model of self-organization", *Journal of theoretical biology*, 120, pp. 371–380.
- Atlan H, Koppel M (1997) 'Self-Organization and Computability', *La Nuova Critica*, vol. 30, n° 2, pp. 5-17.
- Atlan, H. (1999) *La fin du tout génétique*, Paris, Inra.
- Aubin, D. (2008) "The memory of life itself: Bénard's cells and the cinematography of self-organization", *Studies in History and Philosophy of Science Part A*, vol. 39, Issue 3, pp. 359-369.
- Bachelard, G. (1941) *Philosophie du non*, Paris, PUF.
- Bailly, F., Longo, G. (2006) *Mathématiques et sciences de la nature. La singularité physique du vivant*, Paris, Hermann.
- (2007) "Randomness and determination in the interplay between the continuum and the discrete", Special issue: *Mathematical Structures in Computer Science*, vol. 17, n° 2, pp. 289-307.
- (2008) "Extended critical situations", *Journal of Biological Systems*, vol. 16, n° 2, pp. 309-336.
- (2009) "Biological organisation and anti-entropy", *Journal of Biological Systems*, vol. 17, n° 1, pp. 1–34.
- Bailly, F., Longo, G., Montevil, M. (2011) "A 2-dimensional geometry for biological time", *Progress in Biophysics and molecular biology*, 106, pp. 474-484.
- Barabási, A.L., Albert, R. (1999) "Emergence of scaling in random networks" *Science*, 286, pp. 509–512.
- Barbaras, R. (2003) *Vie et intentionnalité*, Paris, Vrin, Paris.
- Barbaras, R. (2008) *Introduction à une phénoménologie de la Vie*, Vrin, Paris.
- Beistegui, M. de (2004) *Truth and Genesis: Philosophy as Differential Ontology*, Indiana University Press.
- Ben Jacob, E., Aharonov, Y., Shapira, Y. (2004) "Bacteria harnessing complexity", *Biofilms*, 1, pp. 239-263.
- Bergé, P., Pommeau, Y., Vidal, Ch. (1997) *L'ordre dans le chaos*, Paris, Hermann.
- Bergson, H. (1896) *Matière et mémoire*, Paris, Alcan.
- (1907) *L'évolution créatrice*, Paris, Alcan.
- (1934) *La pensée et le mouvant*, Paris, Alcan.
- Bertalanffy, L. von (1950) "The theory of open systems in physics and biology", *Science*, vol. 111, pp. 23-9.
- Bitbol, M. (2007) "Ontology, Matter, Emergence", *Phenom Cogn Sci*, 6, pp. 293-307.
- Braudel, F. (1969) *Écrits sur l'histoire*, Paris, Champs Flammarion.
- Bricmont J (1995) "Science of Chaos or Chaos in Science ?", *Physicalia Magazine*, vol. 17, n° 3-4, pp.159-208.
- Brilloui, L. (1959) "La science et la théorie de l'information", Paris, Masson.
- Bunge, M. (2003) *Emergence and convergence*, Toronto: Toronto University Press.
- Bütschli, O. (1892) *Untersuchungen über mikroskopische Schäume und die Struktur des Protoplasmas*. Leipzig, Verlag W. Engelmann.
- Canguilhem, G. (1943) *Le normal et le pathologique*, Paris, PUF.
- Capp, JP. (2012) *Nouveau regard sur le cancer*, Paris, Belin.
- Carosella, E., Pradeu, Th. (2010) *L'identité, la part de l'autre*, Paris, Odile Jacob.
- Cassirer, E. (1977), *Substance et fonction*, Paris, Minuit.

- Collet, P., Eckmann, JP., Lanford O., (1980) "Universal Properties of Maps on an Interval", *Commun. Math. Phys.* 76, pp. 211-254.
- Coquillard, P., Hill, D. (1997) *Modélisation et simulation d'écosystèmes*, Paris, Masson.
- Cornish, A., Cardenas, M.L., Letellier, J.C., Soto-Andrade, J. (2007) "Beyond reductionism: metabolic circularity as a guiding vision for a real biology of systems", *Proteomics* 7, pp. 839-845.
- Coullet, P., Tresser, P. (1978) "Itérations d'endomorphismes et groupes de renormalisation", *Journal de Physique, Colloques*, vol. 39, n° C5, pp. 25-28.
- Cournot, A.A. (1876) *Matérialisme, vitalisme, rationalisme*, Paris, Vrin, 1979.
- Crick, F. (1958) "On Protein Synthesis", *The Symposia of the Society for Experimental Biology* 12, pp. 138-163.
- Danchin, A. (1998) *La Barque de Delphes. Ce que révèle le texte des génomes*, Paris, Éditions Odile Jacob, coll. "Sciences".
- D'arcy, Thomson DW (1917) *On Growth of Form*, Dover.
- Darwin, Ch. (1859) *On the origin of species by means of natural selection*, London, John Murray; tr fr. 6ème édition, Paris, Reinwald, 1882.
- (1868) *The Variation of Animals and Plants under Domestication*, J. Murray, London, 2 volumes; *La variation des animaux et des plantes sous l'effet de la domestication*, trad. fr., Paris, Reinwald, 1868.
- (1862) *On the various contrivances by which British and foreign orchids are fertilized by insects*, I et II, London, John Murray, 1862.
- Davidson, E.H., Erwin, D.H. (2006) "Gene regulatory networks and the evolution of animal body plants", *Science*, 311, pp. 796-800.
- Dawkins, S. (1976) *The selfish Gene*, Oxford University Press.
- Deleuze, G. (1966) *Le Bergsonisme*, Paris, PUF.
- (1983) *L'image-mouvement*, Paris, Minuit.
- (1991) *Qu'est-ce que la philosophie ?* Paris, Minuit.
- Dennett, D. (1995) *Darwin's Dangerous Idea*, New York, A Touchstone Book.
- Devaney, R. L. (1989) *An introduction to chaotic dynamical systems*. Addison-Wesley Publishing Company Advanced Book Program, Redwood City, CA, second edition.
- Dobzhansky, Th., Boesinger, E. (1968) *Essais sur l'évolution*, Paris, Masson.
- Duboule D., Wilkins A. (1998) "The Evolution of Bricolage", *Trends in Genetics*, vol. 14, Issue 2, 54-59.
- Duhem, P. (1906) *La théorie physique son objet et sa structure*, Paris, Chevalier et Rivière.
- (1903) *Évolution de la mécanique*, Paris.
- Dupré, J. (1993) *The disorder of things*, Harvard University Press.
- Eigen, M., Schuster, P. (1979) *The Hypercycle: A Principle of Natural Self-Organization*, Springer-Verlag, New York.
- Einstein, A., Podolsky, P., Rosen, N. (1935) "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Physical Review* vol. 47, pp. 777-780.
- Engel, P. (1989) *La norme du vrai*, Paris, Gallimard.
- Espagnat, d' B. (1997) "Physique et réalité", in: Bitbol et Laugier ed., Editions Frontières.
- Fallen, C. (2012) *L'anomalie créatrice*, Paris, Kimé.
- Farge, E. (2003) "Mechanical induction of twist in the Drosophila foregut/stomodeal primordium" *Current Biology*, 13, pp. 1365-1377.

- Feigenbaum, M. J. (1978). "Quantitative Universality for a Class of Non-Linear Transformations" *J. Stat. Phys.* 19, pp. 25–52.
- Fisher, R. (1936) "Indeterminism and Natural Selection", *Philosophy of Science*, 1, pp. 99–147.
- Fox, Keller E. (1995) *Refiguring Life*, Columbia Un. Pr.
- (2002) *Making Sense of Life*, Harvard University Press; *Expliquer la vie*, Paris, Gallimard, 2004.
- (2007) "The disappearance of function from self-organizing systems", in *Systems biology: biological foundations*, dir. Fred Boogerd, Elsevier, pp. 303–319.
- François, A. (2009) *Bergson, Schopenhauer, Nietzsche, Volonté et réalité*, Paris, PUF.
- Gánti, T. (1975) "Organization of chemical reactions into dividing and metabolizing units: the chemotons", *Biosystems*, 7, pp. 189–195.
- Gayon, J. (1992) *Darwin et l'après Darwin*, Paris, Kimé.
- Giere, R.N. (1999) *Science without Laws*, Chicago: University of Chicago Press.
- Gilbert, S.F., Epel, D. *Écological Developmental Biology: Integrating Epigenetics, Medicine, and Evolution*, Sunderland MA: Sinauer Associates, 2009.
- Gilbert, S.F., McDonald, E., Boyle, N., Buttino, N., Gyi, L., Mai, M., Prakash, N., Robinson J (2010) "Symbiosis as a source of selectable epigenetic variation: taking the heat for the big guy", *Proceeding of Royal Society London, ser. Biol Sci*; 365(1540), pp. 671–8.
- Gilbert, S. et alli (2012) "A symbiotic view of life", *The quaterly Review of Biology*, vol. 87, n°4, pp. 327–341.
- Gödel, K. (1931) "Überformalunentscheidbare Sätze der *Principia mathematica* und verwandter Systeme", *Monatshefte für Mathematik und Physik*, 37, 173–198.
- Gould, S.J. et Lewontin, R. (1979) "The spandrels of San Marco and the panglossian paradigm : a critique of the adaptationist programme", *Proceeding of Royal Society London, ser. B* 205, pp. 581–98.
- Gould, S.J. (1985) *The Flamingo's Smile*, Norton and Company. *Le sourire du flamant rose*, Paris, Seuil, 1988.
- (1989) *Wonderful Life*, Norton and Company; *La vie est belle. Les surprises de l'évolution*, Paris, Seuil, 1991.
- (1996) *Full House*, New York, Harmony Books.
- Heidegger, M. (1927) *Sein und Zeit*; Achtzehnte Auflage, Tübingen: Max Niemeyer Verlag, 2001.
- Guarente, L. (2003) *Ageless Quest*, Cold Spring Harbor Laboratory Press, New York.
- Gros, F. (1986) *Les secrets du gène*, Paris, Odile Jacob.
- Hempel, K. (1965) *Aspects of Scientific Explanation*, The Free Press, N.Y
- Hénon, M. (1976) "A Two-dimensional Mapping with a strange attractor", *Communications in Mathematical Physics*, 50, pp. 69–77.
- Huang, S. et alli (2009) "Non-genetic heterogeneity – a mutation-independent driving force for the somatic evolution of tumors", *Nature Genetics*, vol. 10, pp. 336–342.
- Hull, D. (1980) "Individuality and selection", *Annual Review of Ecology and Systematics*, Vol 11, pp. 311–332.
- Husserl, E. (1907) *Idee der Phenomenologie*. Reed. Walter Biemel (1973) The Hague, Netherlands: Martinus Nijhoff.
- Hume, D. (1739) *A Treatise of Human Nature*, M.J. Norton ed., 2000,

- Ikehata, H. and Alli (2003) "UVA induce C—T mutations at methylated CPG sites", *Mutagenesis*, 18 (6), pp. 511-519.
- Jablonka E., Lamb, M. (2002) The Changing concept of Epigenetics, *N.Y.Acad.Sci*, 981, pp. 82-96.
- (2005) *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*, Cambridge MA: MIT Press.
- Jacob, F. (1970) *Logique du vivant*, Paris, Gallimard.
- (1981) *Le jeu des possibles*, Paris, Fayard.
- Jankélévitch, V. (1931) *Henri Bergson*, Paris, F. Alcan.
- Jenuwein, Th., Allis, D. (2001), Translating the Histone Code, *Science*, 293, pp. 1074-1079.
- (2003) "Suv39h-Mediated Histone H3 Lysine 9 Methylation Directs DNA Methylation to Major Satellite Repeats at Pericentric Heterochromatin", *Current Biology*, vol. 13, 1192–1200.
- Jonas, H. (1966) *The Phenomenon of Life. Toward a Philosophical Biology*, Northwestern University Press.
- Kant, I. (1781) *Kritik der reinen Vernunft*; Paris, Editions Gallimard, Pléiade, 1980.
- Kauffman, S. (1993) *The Origin of Order*, Oxford University Press.
- (2000) *Investigations*, Oxford University Press.
- Kim, J. (1993) *Supervenience and mind*, Cambridge University Press.
- (1998) *Philosophy of mind*, Westview Press.
- (1999) "Making sense of emergence", *Philosophical Studies*, 95, pp. 3–36.
- Kirkwood, T.B. (November 1977), "Evolution of ageing", *Nature*, 270 (5635), pp. 301–4.
- Klein, E. (2007) *Le facteur temps ne sonne jamais deux fois*, Paris, Flammarion.
- Kupiec, J.J., Sonigo, P. (2000) *Ni Dieu ni gènes*, Paris, Seuil.
- Kupiec, J.J. (2009) *L'origine des individus*, Fayard, Paris.
- Lacroix, J., Poët, C., Maehrel, C., Counillon, L. (2004) "A mechanism for the activation of the Na/H exchanger NHE1 by intracellular acidifications and mitogens", *Embo Reports*, 5(1) pp. 91-96.
- Ladrière, J. (1969) "Le théorème de Löwenheim-Skolem", *Cahiers pour l'analyse* 10, pp. 108-130.
- Lapoujade, D. (2010) *Puissances du temps*, Paris, Minuit.
- Laugier, S. (2002) *L'anthropologie logique de Quine*, Paris, Vrin.
- Leibniz, G. (1686) *Discours de métaphysique*, Paris, Vrin.
- Lennox, J.G. (1993) Darwin was a teleologist, *Biology and Philosophy*, 8, pp. 409-421.
- Lesne, A. (2003) *Approches multi-échelles en physique et en biologie*, Habilitation à Diriger les Recherches, Université Paris VI.
- Lesne, A., Victor, J.M. (2006) Chromatin fiber functional organization : some plausible models. *European Physical Journal E* ; 19 : 279-90.
- Lesne, A. et alli (2008) *Physique statistique. Chaos et approches multi-échelles*, Paris, Belin.
- Levin, S.A. (2003) "Complex adaptative systems: exploring the known and the unknown and the unknowable", *Bull. Am. Soc.*, 40 (1): 3-19.
- Lévy-Leblond, J.M. (1996) *Aux contraires*, Paris, Gallimard.
- (2006) *La vitesse et l'ombre*, Paris, Seuil.

- Libet, B. (1999) "How does conscious experience arise ? The neural time factor", *Brain research Bulletin*, vol. 50, n° 5/6, pp. 339-340.
- Lloyd, Morgan (1923) *Emergent Evolution*, Henry Holt and Co.
- Longo, G. (1999) "Mathematical intelligence, infinity and machines: Beyond the Gödelitis", *Journal of Consciousness Studies*, special issue on Cognition, vol. 6, 11-12, 1999.
- (2000) "Cercles vicieux mathématiques et formalisations logiques", *Math.& Sci.hum.* 38^{ème} année, n°152, pp. 5-26.
- (2002) "Reflections on Incompleteness", *Lecture Notes in Computer Science*, vol. 2277 (Callaghan et al. ed.) Springer: 160- 180.
- (2011) *Interfaces de l'incomplétude, pour Les Mathématiques*, Editions du CNRS.
- Luigi, Luisi P. (2003) "Autopoiesis a review and a reappraisal", *Naturwissenschaften* 90:49–59.
- Mach, E. (1883) *Die Mechanik in ihrer Entwicklung historisch kritisch dargestellt*. Leipzig: F. A. Brockhaus.
- McTaggart, J.M.E. (1908) *The unreality of time*, *Mind* 17: 456-473.
- Maine de Biran, P. (1805) *Mémoire sur la décomposition de la pensée*, Paris, Vrin, 1988.
- Malaterre, C. (2011) "Making sense of downward causation in manipulationism: illustrations from cancer research", *History and Philosophy of Life Sciences*; 33(4) : 537-61.
- Margulis, L., Sagan D (1986) *Microcosmos*, Summit Books.
- Margulis, L., Fester R (1991) *Symbiosis as a Source of Evolutionary Innovation: Speciation and Morphogenesis*, MIT Press.
- Margulis, L., Sagan D (2001) The beast with five genomes, *Natural History* 110 :38–41.
- Maynard Smith, J., Szathmary, E. (1999) *The Origins of Life: from the birth of life to the origins of language*, Oxford University Press.
- Mayr, E. (1961) "Cause and Effect in Biology", *Science*, 134: 1501-1505.
- Meillassoux, Q. (2006) *Après la finitude*, Paris, Seuil.
- Merleau Ponty, M. (1979) *Le visible et l'invisible*, Paris, Tel Gallimard.
- Merleau Ponty, M. (1966) *Sens et non sens*, Paris, Nagel.
- Meyerson, E. (1908) *Identité et réalité*, Paris, Alcan.
- Miquel, P.A. (2000) *Comment penser le désordre ?* Paris, Fayard.
- (2004) "Considérations épistémologiques sur les cercles vicieux épigénétiques du vieillissement", *Journal de la Société Française de Biologie*, 198 (3).
- Miquel, P.A., Soto A, Sonnenschein C (2008) "Physicalism and downward causation in developmental and cancer biology", *Acta Biotheoretica*, 56, N° 4, pp. 257-274.
- Miquel, P.A., Perovic S (2010) "On genes action and reciprocal causation", *Foundations of Science*, Vol 16, 2011, pp. 31-46.
- Monod, J. (1970) *Le hasard et la nécessité*, Paris, Seuil.
- Montebello, P. (2003) *L'autre métaphysique : essai sur la philosophie de la nature*, Ravaisson, Tarde, Nietzsche et Bergson, Paris, Desclée de Brouwer, " Philosophie ".
- Montebello, P. (2006) *Nature et subjectivité*, Grenoble, Million.
- Montévil, M. (2011) *Temps biologique et transitions critiques étendues*, Thèse de Doctorat es Science.
- Morange, M. (1998) *La part des gènes*, Paris, O. Jacob.
- (2003) *Histoire de la biologie moléculaire*, Paris, La découverte.

- (2003) *La vie expliquée ?* Paris, Odile Jacob.
- (2011) *La vie, l'évolution, l'histoire*, Paris, Odile Jacob.
- Moreno, A., Etxeberria, A. & Umerez, J. (2008), "The autonomy of biological individuals and artificial models", *Biosystems*, 91, pp. 309-319.
- Mossio, M., Longo, G., Stewart, J. (2009) Computability of closure to efficient causation, *J. of Theoretical Biology*, 257, 3. pp. 489-498.
- Mossio, M., Moreno, A. (2010) "Organisational closure in biological organisms", *History and philosophy of life sciences*, vol. 32, pp. 269-288.
- Nagel, E. (1961) *The Structure of Science: Problems in the logic of Scientific Explanation*, N.Y.: Harcourt, Brace & World.
- Nagel, E. (1997) *Le théorème de Gödel*, presentation, Paris, Seuil, Point Sciences.
- Nicolis, G., Prigogine I (1989) *Exploring Complexity*, New York, Freeman.
- Nan, X. et al (1998) "Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex", *Nature*, 393 : 386-389.
- Noble, D. (2002) "Modelling the heart: insights, failures and progress", *BioEssays* 24 : 1155-1163
- Noble, D. (2006) *The Music of Life. Biology Beyond the Genome*, London, Hardback.
- Noble, D. (2008) "Genes and Causation", *Phil. Trans. R. Soc. A* , 366 : 3001-3015.
- Perovic, S. (2007) "The limitation of Kim's reductive physicalism", *Acta Biotheoretica* 55, pp. 243-267.
- Poincaré, H. (1902) *La science et l'hypothèse*, Paris, Flammarion.
- Popper, Karl R. (1935) *Logik der Forschung*, tr.fr. Nicole Thyssen-Rutten, Paris, Payot, 1973.
- Pouvreau, D. (2013) *Une histoire de la "systémologie générale" de Ludwig von Bertalanffy*, Thèse de doctorat.
- Pradeu, T., Carosella E.D. (2006), "The Self Model and the Conception of Biological Identity in Immunology", *Biology and Philosophy* 21(2), pp. 235-252.
- Pradeu, T. (2010), What is an organism? An immunological answer, *History and Philosophy of the Life Sciences* 32, pp. 247-268.
- Pradeu, T. (2011) "A mixed self : the role of symbiosis in development", *Biological Theory* 6, pp. 80-88.
- Prigogine, I. (1969) Symmetry Breaking Instabilities in Biological Systems, *Nature* 223, 913-916.
- Prigogine, I., Stengers, I. (1979) *La Nouvelle alliance*, Paris, Gallimard.
- Prigogine, I. Nicolis, G. (1989) *Exploring complexity*, W H Freeman & Co Ltd.
- Prusiner, S. B. (1991) Molecular biology of prion diseases, *Science* 252 (5012), pp. 1515-1522.
- Putnam, H. (1988) *Representation and Reality*, Reed Business Information, Inc.
- (1990) *Realism as a Human Face*, Harvard Un. Pr.
- Quine WV (1969) *Ontological Relativity and Other Essays*, New York, Columbia University. Press, tr. Fr. Paris, Aubier, Montaigne, 1977.
- (1990) *Pursuit of Truth*, Harvard University Press.
- Rashevsky, N. (1938) *Mathematical Biophysics: Physico-Mathematical Foundations of Biology*, Chicago. University of Chicago Press.
- Radman, M. (1999 a) "Enzymes of Evolutionary Change", *Nature*, 401: 866-69.
- Radman, M. et alli (1999b) "Evolution of Evolvability", *N.Y. Ac. Of Sc*, 870 : 146-155.

- Radman, M. (2003) "Entretien avec Miroslav Radman" in Marie-Christine Maurel et Paul-Antoine Miquel *Nouveaux débats sur le vivant*, Paris, Kimé, pp. 106-115.
- Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., Tyagi, S. (2006) "Stochastic mRNA synthesis in mammalian cells", *PLoS Biol* 4(10): e309. doi:10.1371/journal.pbio.0040309.
- Rassoulzadegan, M. et alli (2006) "RNA-mediated non-Mendelian inheritance of an epigenetic change in the mouse", *Nature* 431, pp. 469–474.
- Rausser, C.L., Tierney, J.J., Gunion, S.M., Covarrubias, G.M., Mueller, L.D., Rose, M.R. (2006) "Evolution of late-life fecundity in *Drosophila melanogaster*", *Journal of Evolutionary Biology*, 19 : 289-301.
- Riquier, C. (2009) *Archéologie de Bergson*, coll. "Epiméthée", Paris, PUF.
- Robert, L. (1995) *Le vieillissement, faits et théories*, Paris, Flammarion, domino, 1995.
- Robert, L. (1997), *Les Mécanismes épigénétiques du vieillissement*, C.R. Société de Biologie, 191, 247-52.
- Ronchi, Rocco, *Bergson filosofo dell'interpretazione*, Marietti, 2000.
- Rose, M.R. et alli (2006) "A revolution for aging research", *Biogerontology* (2006) 7, pp. 269–277.
- Rose, M., Oakley, T. (2007) "The new biology, beyond the modern synthesis", *Biology direct*, 2, 30, doi:10.1186/1745-6150-2-30.
- Rosen, R. (1972) "Some relational cell models: the metabolism-repair system", in: Rosen R. ed., *Foundations of Mathematical Biology*, New York, Academic Press.
- Rosenberg, A., Kaplan, D.M. (2005) "How to reconcile Physicalism and Antireductionism about Biology", *Philosophy of Science*, 72, pp. 43-68
- Rosenberg, A. (2006) "Is epigenetic inheritance a counterexample to the central dogma?", *History and Philosophy of the Life Sciences*, 28, pp. 549–566.
- Russell, B. (1914) *Our Knowledge of the External World as a Field for Scientific Method in Philosophy*, London, Routledge, 1914.
- Sapp, J. (2009) *The New Foundations of Evolution: On the Tree of Life*. New York: Oxford University Press.
- Schoffeniels, A. (1973) *L'anti-hasard*, Paris, Gauthier-Villars.
- Schrödinger, E. (1944) *What is life?* Cambridge University Press.
- Simondon, G. (1964) *L'individu et sa genèse physico-biologique*, PUF, Paris.
- Smolin, Lee (2006), *The trouble with physics*, Houghton Mifflin.
- Sober, E. (1984), *The Nature of Selection: Evolutionary Theory in Philosophical Focus*, Cambridge, MA: MIT Press.
- Solé, R.V., Munteanu A, Rodriguez-Caso, C, Macía J (2007) Synthetic protocell biology: from reproduction to computation. *Philosophical Transactions of the Royal Society B*, 362, pp. 1727-1739.
- Sonnenschein, C., Soto, A. (1999) *The Society of Cells*, Bios Scientific Publishers, Oxford.
- Soto, A., Sonnenschein, C. (2004) "Emergentism as a default: Cancer as a problem of tissue organization", *J. Biosci*, 30, pp. 103–118.
- Soto, A., Sonnenschein, C., Krause, S., Maffini (2008) "A novel 3 D in vitro culture to study stroma/epithelial interactions in the mammary gland", *Tissue Engineering Part C: Methods*, 14(3), pp. 261-271.
- Soto, A.M. and Sonnenschein, C. (2009) "The tissue organization field theory of cancer: A testable replacement for the somatic mutation theory", *BioEssays* 33, p. 332.

- Spencer, H. (1864) *Principles of Biology*, London, Williams and Norgate.
- Stengers, I. (1997) *Cosmopolitiques VI : La vie et l'artifice : visages de l'émergence*, Paris, La découverte.
- Stephan, A. (1999) "Varieties of Emergentism", *Evolution and Cognition*, vol. 5, n°1, 54.
- Strahl, B., Allis, C. (2000) "The language of covalent histone modifications", *Nature* 403 (6765) : 41–5.
- Suppe, F. (1989), *The Semantic Conception of Theories and Scientific Realism*, Urbana: University of Illinois Press.
- Uexküll, J. Von (1934) *Streifzüge durch die Umwelten von Tieren und Menschen: Ein Bilderbuch unsichtbarer Welten* (Sammlung: Verständliche Wissenschaft, Bd. 21) Berlin: J. Springer (mit Kriszat G.) ; tr. Fr. *Milieu animal et milieu humain*, Paris, Payot-Rivages, 2010.
- Taniguchi, Y., et al., (2010) "Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells", *Science* 329: 533-38.
- Turing, A. (1937) "On computable numbers, with an application to the Entscheidungs Problem", *Proc. London Math. Soc.*, 2^e série, vol. 42 : 230-265.
- Van Fraassen, B. (1990) *Laws and symmetry*, Oxford University Press.
- (2004) "Science as representation: flouting the criteria", *Philosophy of Science* 71, pp. 794–804.
- Varela, F., Maturana, H., Uribe, R. (1974) "Autopoiesis: the organization of living systems, its characterization and a model", *Biosystems* 5, pp. 187-196.
- Varela, F. (1980) *Principles of Biological Autonomy*, North Holland/Elsevier, New York; *Autonomie et connaissance*, Paris, Seuil, 1989.
- Varenne, F. (2007) *Du modèle à la simulation informatique*, Paris, Vrin, collection Mathesis.
- Vidal, F., Miquel, P.A. (2008) "Du concept de gène au concept d'épigénétique en biologie", in: *Biologie du XXIème Siècle, Evolution des concepts fondateurs*, Bruxelles, De Boeck: 219-243.
- Virchow, R.C. (1858) *Die Cellularpathologie in ihrer Begründung auf physiologische und pathologische Gewebelehre: 20 Vorlesungen*, Berlin, Hirschwald.
- Volterra, V. (1928) "Variations and fluctuations of the number of individuals in animal species living together", *J. Cons. int. Explor. Mer* 3(1) : 3-51.
- Von Kiedrovski, G. et alli (1994) "Self-replication of Contemporary nucleotide oligomers", *Nature*, 369, pp. 221-24.
- Von Neumann, J. (1961) "The General and Logical Theory of Automata", in: Taub ed., *John von Neumann: Collected Works. Volume V: Design of Computers, Theory of Automata and Numerical Analysis*, Pergamon Press, Oxford, 9, pp. 288-328.
- Waddington, C.H. (1968, 1969, 1970, 1972) ed., *Towards a theoretical Biology*, Edinburgh University Press, Edinburgh, 4 volumes.
- Weismann, A. (1892) *Der Keimplasma, eine Theorie der Vererbung*, Fischer, Jena.
- West-Eberard, M.J. (2003) *Developmental Plasticity and Evolution*, Oxford University Press.
- Westerhoff, H., et alli, (2009) *Systems biology: the elements and principles of life*. FEBS Letters 583, 3882e3890.
- Wilson, K. G. (1971) "Renormalization group and critical phenomena", *Physical Review B*, 4, pp. 3174- 3183.

- Wimsatt, W. (2007) *Re-engineering philosophy for limited beings*, Cambridge: Harvard University Press.
- Wittgenstein, L. (1953) *Philosophisches Untersuchungen*, Blackwell Publishers.
- Wittgenstein, L. (1921) *Tractatus Logico-philosophicus*, Ann. Nat. Phi.
- Worms, F. (2004) *Bergson ou les deux sens de la vie*, PUF, Quadrige, Paris.
- Wright, S. (1932) The Roles of Mutation, Inbreeding, Crossbreeding, and Selection. In: D.F. Jones (ed.) *Evolution. Proceedings of the Sixth International Congress of Genetics*: 356-366.
- Wright, S. (1977) *Evolution and the Genetics of Population*, T3. Univ. Chicago Press, Chicago.
- Wright, S. (1980) “Genic and Organismic Selection”, *Evolution*, 34: 825-843.
- Wu, C. and Morris, J.R. (2001) “Genes, genetics and epigenetics: a correspondence”, *Science* 293, pp. 1103–1105.
- Zeh, H.D. (1970) “On the Interpretation of Measurement in Quantum Theory”, *Foundation of Physics*, vol. 1, pp. 69–7.

Causes and Entities in Biology*

1. Biological individuals and causes

A number of different criteria of biological individuality have been proposed, ranging from reproductive capacity, as a very general but biologically sound criterion, to those that emphasize various structural features.¹ Following the latter, organism has been a favourite candidate. But other levels have been treated likewise. For instance, biologists and philosophers that focus on molecular biological processes find it appropriate to treat molecules that carry developmental information, such as DNA, RNA, and recently various elements of proteins, as the most basic individual elements of life.²

Irrespective of the choice of the level at which biological individuals are defined, it seems that the inherent nature of biological explanations renders any concept of “basic entity” very tenuous. In a *prosaic* sense, living entities do not have sharp boundaries even when viewed with a naked eye. Yet, more careful considerations also demonstrate that biological entities do indeed have vague boundaries. It quickly becomes an arbitrary choice of where to draw the line in delineating them. There seems to be no principled way of deciding to draw one line rather than another. Thus, there is no unique determination of building-block entities. For instance, the more we know about genes, the more the *relevance of genes’ sequences* seems dependent on the choice of the way we sequence them and on the goals the sequencing serves.³ Also, it has become apparent that the

* This work was supported by the project “Dynamic Systems in nature and society: philosophical and empirical aspects” (evidence #179041) financed by the Ministry of Education, science and technological development of the Republic of Serbia.

¹ J. Wilson, *Biological Individuality: The Identity and Persistence of Living Entities*, New York, Cambridge University Press 1999. J. Wilson, “Ontological Butchery: Organism Concepts and Biological Generalizations”, *Philosophy of Science* 67, 2000, pp. 301-311.

² J. Maynard Smith, “The Concept of Information in Biology”, *Philosophy of Science* 67, 2000, pp. 177-194. R. Dawkins, *The Selfish Gene*, Oxford, Oxford University Press, 2nd edition 1989 (1st edition 1976).

³ P.E. Griffiths, K. Stotz, “Gene”, in: D.L. Hull and M. Ruse eds., *The Cambridge Companion to Philosophy of Biology*, Cambridge, Cambridge University Press, 2007.

process of *coding* itself is not uniquely defined, because it is dependent on the context of the research.⁴

Generally speaking, it does not seem to be in the spirit of biological theory to ask where a single molecule, a molecular complex, or a living entity ends, and whether it can be treated as a causally functional basic unit; parts/whole distinctions are not uniquely differentiated but are entangled functional units. But once we start thinking of *causal powers of, and interactions among* biological individuals, the failure of metaphysical views of life predicated on the existence of basic individuals becomes even more transparent.

Thus, the so-called *endosymbiosis*, an account of living cells within other living cells, regards different forms of life as mutually dependent. It is hard to see what exactly one could achieve were one to seek a basic sort of symbiot. It is much more advantageous to talk about interacting *living entities* at different levels in such cases. For instance, the microorganisms that exhibit continual lateral transfer of genes in complex cell communities are best characterized by their *communal genome*, which makes the distinction between elementary microorganisms and a community of such organisms rather tenuous. The so-called *superorganisms* of such microbiological communities can be hardly treated as simple aggregates of individual microorganisms when they are involved in the *communal metabolic process* as well.⁵ The subvenient/supervient, or their variations distinctions, seem to be missing the target if employed in such cases.

This is also, and perhaps more so, true of viruses that inhabit cells. The so-called capsid-encoding organisms can be treated as basic in the sense that they generate protein programs, but bacteria provide the living environment for them, where their parts and metabolism are entangled with viruses.⁶

The distinction between the organism and living cells — as their basic components — has also been called into question. Thus, the cells in plants are understood both as its active building blocks but also as mere compartments that originate in the walls of the

⁴ Brosius J., “The Fragmented Gene”, *Annals of the New York Academy of Sciences*, vol. 1178, 2009, pp. 186-193.

⁵ J. Dupre, A. O. O’Malley, “Metagenomics and biological ontology”, *Studies in History and Philosophy of Biological and Biomedical Sciences*, 38, 2007, pp. 834-846. See also D. S. Wilson and Elliott Sober, 1989, “Reviving the Superorganism”, *Journal of Theoretical Biology*, vol. 136, Wilson D. S. and Elliott Sober, “Reviving the Superorganism”, *Journal of Theoretical Biology*, vol. 136, 1989, pp. 337-356.

⁶ P. Forterre, “Giant Viruses: Conflicts in Revisiting the Virus Concept”, *Intervirology*, vol. 53, 2010, pp. 362-378.

protoplast existing at the level of the organism. Thus cells and organism are best understood as interacting components rather than as different levels of a bottom-up hierarchical organization.⁷ Similarly, in explanations of circadian cycles the cell is treated both as a locus of biochemical processes and as their object.⁸

Even at the level of *the cell membrane*, the entangled levels play a crucial explanatory role. Thus, the so-called lipid rafts, intricately extended molecules of fat in the membrane tissue, can be properly characterized only as *both* supervenient and subvenient — i.e. as both the carrier and the receiver of causal powers that resist characterization in terms of basic level causal instantiations. The membrane protein complexes “carry out signal transduction — as opposed to these same components partitioning into lipid rafts to convey signal”.⁹ As the flow of membrane proteins and lactose are mutually regulated¹⁰, the self-regulation of membranes is better expressed in the concept of interactive layers, rather than hierarchically ordered levels.

In morphogenetic processes we have gene expression that gives rough guidelines for the processes to take place, in combination with the conditional responses of tissues and responses to external forces (i.e., chemical cell communication), as well as the physics of excitable media controlling the flow of fluids. Each molecular level, and each part of the morphogenetic system, acts in a reciprocal manner with other levels.

2. Integrative biological explanations

Prima facie, there seems to be a convincing rejection of the conclusion that one cannot regard certain kinds of biological individuals and causes as basic. This is because they cannot be unequivocally defined in light of actual biological explanations. The response states that particular kinds of *molecules* can be regarded as basic biological

⁷ D. Kaplan, “The Relationship of Cells to Organisms in Plants: The Problem and Implications of Organismal Perspective”, *International Journal of Plant Sciences*, vol. 153, 1992, pp. 28-37.

⁸ Bechtel W., “The cell: locus or object of inquiry?”, *Studies in History and Philosophy of Biological and Biomedical Sciences*, vol. 41, 2010, pp. 172-182.

⁹ Jacobson K., Mouritsen O. G. and Anderson R. G.W., “Lipid rafts: at a crossroad between cell biology and physics”, *Nature Cell Biology*, vol. 9, n° 1, p. 10.

¹⁰ M. Morange, “Articulating Different Modes of Explanation in Biology: The Present Boundary in Biological Research”, Mapping the Future of Biology, *Boston Studies in Philosophy of Science*, vol. 266, 2009, pp. 15-26.

individuals, just as their interactions can be considered fundamental. This constitutes a fixed hierarchy that starts from such molecules as basic, and ends with the cells and individual organisms that they form. This is the crux of a variety of reductionist positions, with regard to inter-level relations in biological systems.

A typical example of a biological explanation that unequivocally conforms to such strong reductionism is an early understanding of the function of DNA. An optimistic version of reductionism states that such a hierarchy can be derived from current biological explanations or their appropriate modifications.¹¹ But this version of reductionism, as well as all other versions, is committed to the view that such derivation is obtainable in principle. And this view applies to both synchronic and diachronic aspects of biological systems — i.e. it concerns both biological constituents and their evolution.

Note that a simple *movement from mereological to causal reduction* underlies reductionist varieties (let us call it reductionist entailment - RE): the physicalist presupposition that a biological whole is reducible to its elementary parts, as, allegedly, is any other whole in the physical world, implies that the causal powers of the whole can be realized only through the causal powers of such parts. What the whole can do, and what it does do, is reducible to what its elementary parts are doing. The colloquial phrase “It’s all molecules and their interactions” is intended to capture this — for example, what a cell can do is reducible to what its basic molecular components can do in mutual interactions.

The much discussed *Causal/Explanatory Exclusion Argument*¹² has motivated, to a great extent, the metaphysical bases of the reductive physicalist stance, showing the basic incoherence of various notions of downward causation that do not conform to the hierarchical account of causal powers. The much discussed argument states that arguing in favour of downward causal influence leads to a destructive dilemma, whereby accepting either the overdetermination (by causal instances at multiple levels) of an effect — which undermines the very foundation of causal explanation on the counterfactual

¹¹ A. Rosenberg and D. M. Kaplan, “How to Reconcile Physicalism and Antireductionism about Biology”, *Philosophy of Science*, vol. 72, 2005, pp. 43-68. S. Perovic and P-A. Miquel, “On Gene’s Action and Reciprocal Causation”, *Foundations of Science*, vol. 16, 2011, pp. 31-46.

¹² J. Kim, *Supervenience and Mind*, Cambridge, Cambridge University Press, 1993. J. Kim, *Mind in a Physical World*, Cambridge Mass., MIT Press, 1998. J. Kim, *Physicalism or Something Near Enough*, Princeton and Oxford, Princeton University Press, 2005.

analysis of the case — or the perils of dualism.¹³ In a nutshell, the argument states that every physical effect can have only one sufficient physical cause, which excludes the possibility of macro-properties realising causal powers; the realisers are always at the micro-level. In one form or another it underlies diverse reductionist positions. For instance, Rosenberg and Kaplan state: “Were the biological facts at t1 to cause or explain the physical facts at t2, they would compete with the explanation of these facts by appeal to the physical facts at t1 or overdetermine the occurrence of the physical facts at t2. [...] Neither alternative is acceptable to the physicalist.”¹⁴ It is not clear whether the Exclusion Argument supports RE, or vice versa, but typically those who criticise the argument do not challenge RE which may be an indication that the latter has metaphysical primacy.

In any case, given the presupposition of mereological reduction and the acceptance of RE, all anti-reductionist accounts postulate biological individuals and causes at levels higher than the molecular, which implicitly implies, or explicitly appeals to downward causation in one form or another. Yet these accounts turn out to be incoherent because of the failure of causal reduction. This constitutes an “ignorance charge” against anti-reductionism in any substantial form: biological explanations appeal to downward causation only in ignorance of the details, because basic entities, their properties and interactions, are not analysed in sufficient depth. Thus, such explanations, although indispensable and useful in some contexts, lack theoretical clarity because they appeal to useful — but essentially metaphysically incoherent — shortcuts. They involve a metaphysically incoherent concept of downward causation and thus cannot provide a coherent concept of biological individuals either.

Some philosophers have responded by attacking the premise of the reductionist argument, concerning biological individuals as reducible to mereological wholes. They take an essentially metaphysical stance concerning coincident entities: mereological wholes coincide with, but are not identical to, living organisms.¹⁵

Another strategy, closer to biological explanations than to the metaphysical premise, is to attack the “ignorance charge”. Thus, the claim is that reductionists neglect

¹³ J. Kim, *Mind in a Physical World*, Cambridge Mass., MIT Press, 1998. S. Perovic, “The Limitations of Kim’s reductive Physicalism in Accounting for Living Systems and an Alternative Nonreductionist Ontology”, *Acta Biotheoretica*, vol. 55, 2007, pp. 243-267.

¹⁴ A. Rosenberg and D. M. Kaplan, “How to Reconcile Physicalism and Antireductionism about Biology”, *Philosophy of Science*, vol. 72, 2005, p. 46-n6.

¹⁵ J. K. Crane, “Biological-mereological Coincidence”, *Philosophical Studies*, vol. 161, n° 2, 2012, pp. 309-325.

the fact that the deeper and more detailed the explanation of biological phenomena, the more entrenched they are in causal reciprocity of levels and the individuals that define them.¹⁶ Such a strategy does not aim to destroy the argument, but rather intends to discourage reductionism by stripping it of its last evidence-based claim — that which concerns the level of molecular biology. The properties that reductionists wish to deem basic physical properties become even more extended in the biological context. Thus, even detailed explanations of supposedly basic molecular levels do not demonstrate a fixed hierarchy. Rather *causal reciprocity* is invoked through and through in biology: biological explanations integrate levels in a way that posits causal flow in multiple directions (i.e., causal reciprocity among levels), and describe entities at multiple levels that realise such flows as basic, in a substantial sense.

Thus we find entangled hierarchies of *molecular layers*. Sure, “it’s all molecules”, but which parts of molecules and which aggregates are basic is a question that does not have any straightforward answer. Actually, at every level the elements of a structured organism proteins and amino-acids tinker with the molecular processes responsible for the formation of that very same organism. Analysis of mere hydrogen bonds that tie the molecular building blocks of DNA and proteins together has had to be amended after explanations that concerned a particular protein structure and position. This structure extends into larger molecular complexes, including cellular and tissue structure — if one requires a more detailed explanation. Thus, *molecular bonds* <-> *molecular bonds* interactions are increasingly merged with *molecular aggregates* <-> *molecular bonds* interactions as the explanations provide more details of these cases.

For instance, sticking to the idea that DNA sequences are basic interacting elements does not suffice, as the concept of 2D sequence information as the basic level is elusive. It turns out that interactions with the proteins determines the position of the sequence, which is crucial for the sequence expression.¹⁷ Thus, there are 2D and 3D structures of DNA sequences, and accordingly, there are two types of the same base-pair when we look at it in terms of its geometry (i.e., its hydrogen bonding patterns). Proteins *shift the equilibrium* between two types; although they have evolved to *recognize* only one of them (namely, the 2D structure). Whether a protein will interact with DNA will depend on this (i.e., on whether the protein shifted the equilibrium to the state it can “read”). Thus, the

¹⁶ S. Perovic and P-A. Miquel, op.cit.

¹⁷ B. Honig and R. Rohs, “Flipping Watson and Crick”, *Nature*, vol. 470, 2011, pp. 472-473.

base-pair introduces particular conformational topological shifts in the processes along with the regular chemical structure; where in effect we have *genome* $\leftrightarrow$ *cellular structure interactions*, rather than just a molecular layer as the basic theoretical posit.

Similarly, much recent research has demonstrated that interactions between DNA molecules and chromatin, initially thought to be an inert scaffolding on which DNA hangs, should be understood as entangled micro- and macro-properties of molecules that conform to each other.¹⁸

Finally, the basic mechanism of gene regulations consists of the trans-regulation and the cis-regulation of gene expression. Now, the DNA (i.e., its sequences) codes for transcription regulators (amino-acids) but elements of these trans-regulators bind cis-regulators, namely, the sequences of DNA. We are dealing here with a continuous molecular interaction loop, where the attempt to employ a template of a basic-molecular-level becomes obsolete. There is another crucial DNA-protein loop, where the so-called Tet proteins act as mediators in DNA methylation — the process of adding or leaving out methyl groups on the genome sequences — the process which (not the sequences alone) determine the structure of tissues.

Thus, we find only ubiquitous molecular units cross-determined by multiple levels of molecular complexes/aggregates, rather than anything like basic molecular units acting in unidirectional manner. Why should one hope to divide such molecular systems into unidirectional hierarchies if we take their functional properties into account? Where does a single molecule end, that can be treated as a causally functional basic unit? Shouldn't they be treated properly as molecular systems with multiple entangled divisions, divided into functional units, and as such constituting the basis of explanatory power?

How exactly would a final *molecular bonds* $\leftrightarrow$ *molecular bonds* explanation look like is not clear, and in any given actual case even the simplest molecular explanation will not suffice; molecular aggregates of various structure and complexity are part of the causal story.

We have something like the following overall map of interactions:

Genes $\leftrightarrow$ proteins $\leftrightarrow$ cell structures $\leftrightarrow$ organism

¹⁸ M.H. Kagey et al., “Mediator and cohesion connect gene expression and chromatin architecture”, *Nature* 467, vol. 2010, pp. 430-435; A. Lesne and J.-M. Victor, “Chromatin fiber functional organization: Some plausible models”, *The European Physical Journal E*, vol. 19, 2006, pp. 279-290. See also Perovic and Miquel 2011.

where all the elements are basic theoretical points in one context or another, employed with reciprocal causal powers. The bonds isolated for the purpose of analysis are always enabled by particular properties of larger complexes, and vice versa; they are part of such enabling complexes in other contexts. Biological explanations often invoke these different levels and their capacities to explain phenomena that depends on such integration. All of the examples we have discussed are this same kind of *integrative explanation*. Not all biological explanations are such; some deal with quite narrow phenomena and need to invoke only the elements of a particular level. *But the more elaborate and deeper investigated the phenomenon, the more integrative its explanation becomes, as we have seen in the case of microorganisms or molecular processes of life.*

Thus, there is a clear failure in any analogy with machines with fixed parts, as we are dealing with shifting hierarchies in integrative biological explanations.¹⁹ Organisms are entities that can be divided meaningfully in infinitely many ways and the priorities in such divisions are determined by narrow theoretical goals.²⁰

Irrespective of the outcome of the metaphysical debate on coincident entities, it is clear that mereological reduction is irrelevant and seems to be an impediment to biological theory and explanations if pursued hastily and in a general manner.²¹ It is not important whether employing the entailment from mereological to causal reductionism fails, or if employment of the presupposition of mereological reduction fails because both mereological and causal reduction fail independently in integrative biological explanations.

3. Instrumentalism and other instrumentalist-minded alternatives to reductionism

Understandably, reductionism has lost much of its appeal in the last few decades due to substantial insights in the life sciences that do not square with its assumptions. This

¹⁹ W. Wimsatt, "Emergence as non-aggregativity and the biases of reductionism", *Foundations of Science*, vol. 3, 2000, pp. 269-297.

²⁰ S. A. Kauffman, "Articulation of Parts Explanation in Biology and the Rational Search for Them", *PSA Proceedings*, 1970, pp. 257-272.

²¹ The issue of natural selection at multiple levels is interrelated with the problem we are dealing with here. Yet this does not come up — at least explicitly — to a great extent, because we are focusing on synchronic aspects of interactions across levels and constitutive biological explanations. Such aspects are closely related to the problems and explanations in molecular biology and in related subfields of biology.

has resulted in reconsiderations or simple ignorance of the reductionist view, as well as seeking for alternatives. A shift away from reductionism has produced various positions. Instrumentalism has emerged as a prominent alternative in various guises. The authors defending such positions share a view that biological explanations do not aim at metaphysical coherence regarding entities and causes — as it is defined by reductionist metaphysics — and that any subsequent attempt to derive ontology (reductionist or anti-reductionist) from biological explanations is elusive. But, the argument goes, this is not a difficulty that has to be overcome, because *biological explanations are not inference-based but rather manipulation-based explanations*. These concern the ways in which living beings can be successfully manipulated, and any theory emerging from such activity is of secondary importance. Moreover, there may even be no fundamental theory to start with. Thus, Waters (2008b, 1) suggests: “It is not the fundamental theory of molecular biology, if indeed such a theory exists, that revolutionized genetics and transformed biological sciences. What changed biology so dramatically was a retooling of the investigative strategies used in genetics.”²² Thus, we see the basic theoretical elements such as genes are primarily “used as levers to manipulate and investigate a wide variety of biological processes.”²³ Investigative utility of genetics has been grossly neglected by philosophers: the success of DNA-based research is not a fundamental theory of DNA, which has only a secondary importance.²⁴ Thus “the key for understanding what DNA did for biology is to stop being distracted by theoretical sideshows and focus on the main event.”²⁵

Another suggestion to focus on the methodological aspect of biological explanations and thus to avoid inference-based understanding of them comes from Sitte.²⁶ He believes that there is no real conflict between seemingly disparate causal accounts concerning plant structure and the processes that we have mentioned above, *if positions are not overstated*. They simply reflect different methodological approaches as starting positions. Thus, the analytical and experimental approach of the cellularists, or the

²² K. Waters, “Beyond theoretical reduction and layer-cake anti-reduction.” *Phil-sci Archive*, 2008, p. 1, <http://philsci-archive.pitt.edu/3834/>.

²³ *Ibid.*, p. 23.

²⁴ *Ibid.*, p. 22.

²⁵ *Ibid.* p. 23. See also Waters K., “How Practical Know-How Contextualizes Theoretical Knowledge: Exporting Causal Knowledge From Laboratory to Nature”, *Philosophy of Science*, vol. 75, 2008, pp. 707-719.

²⁶ P. Sitte, “A Modern Concept of the “Cell Theory”: A perspective on competing hypotheses structure”, *International Journal of Plant Sciences*, vol. 153, 1992, pp. 1-6.

morphological and embryological approach of the proponents of the organismal viewpoint are at odds only methodologically. One is justified to make stronger conclusions from this only if one is committed to inference-based view of explanations in biology, which is not necessary.

Similarly, Craver argues that the best framework for capturing the structure of explanations in recent biological studies of neurons is a *pragmatics of levels* that is based on the manipulability explanations.²⁷ The relationship of manipulability should in fact replace the relationship of derivability for constitutive explanations. The traditional understanding of explanation infers, from the study of a relevant mechanism or an inter-level structure, a phenomenon that is characterised either by covering laws that relate its levels, or by a causal-mechanical account. Biological explanations, however, focus on successful manipulation, i.e., knowledge of changes under variety of conditions that result in the production of regular changes.

The claim common to these accounts is that one should not agonize over the fact that biological explanations are typically a patchwork of disparate approaches and models, which together fail to provide a coherent basis for inferring ontological accounts of entities and causes, at least as long as they generate successful manipulations. Now, what these various instrumentalist-minded and reductionist accounts have in common is doubt *that integrative biological explanations can tell us much about the nature of living systems, because of the supposedly incoherent foundations of a desired ontology that contradict RE*, although it is true that they disagree on the prospects and the nature of biological enterprise.

The instrumentalist conclusion is too hasty: biological explanations may both be manipulation devices and yet may often entail coherent ontological frameworks. A problem for reductionist metaphysics and instrumentalism alike is the possibility that integration of levels and their causal entanglement is not only methodological but also explanatory, in the inferential sense.

Instrumentalism about biology is also generally a philosophically suspect position. Why should we subscribe to instrumentalism in biology and not in physics? In physics we have arguably less convincing reasons to become instrumentalists — and it is not clear that we have reasons analogous to those in biology at all. Does this discrepancy between

²⁷ C. Craver, “Constitutive Explanatory Relevance”, *Journal of Philosophical Research*, vol. 32, 2007, pp. 3-20.

our understanding of physics and biology tell us that we cannot have a coherent philosophical view of the nature of science? If not, why should we model it on biology rather than physics?

There are alternatives to reductionism that try to bypass the reductionist commitment to RE. Similarly to Sitte, Kaplan argues that the opposite causal understanding of the constitution and development of the organism imply opposite methodological principles as well.²⁸ But, instead of cautioning against the “overstating” of either position, he argues that one can successfully resolve it: the way in which form and cell lineage are related will answer the question about what is the cause and what is the effect. Thus, if cell division patterns are not the genesis of plant form, the view he favors – based on the existing studies — is that form is (a causal) constraint on cell divisions. It is indeed possible that we might be able to decide, based on the best available evidence, whether an organismal or reductionist (cellularist) view is adequate. The problem we have pointed out, however, is that in many important cases the explanatory power of biological explanations depends on integrating these disparate approaches. Moreover, a possible empirical confirmation of the organismal account would leave it open to the reductionist ‘ignorance charge’ that presumes, in principle, the superiority of the cellularist account.

A similarly piecemeal approach might favor the reductionist approach in some, or even most cases. Love argues in favor of such *contextualization of reduction in biology*.²⁹ But note that it effectively eliminates the ontological basis of reduction by avoiding to posit a molecular level as fundamental, irrespective of the context of the case, and successfully bypassing any meddling with the problem of incoherence of causal reciprocity. Much like the instrumentalist-minded approaches it cannot avoid defining individuals and causes in terms of the aims of manipulation and a very particular scope (context) of phenomena, to which reductionist methodology is then applied.

Morange makes a stronger and a more general claim against the holistic account, from a similar point of view of contextualized reductionism: “The most active works pursued today in systems biology does not seek to replace the molecular description by a holistic one, but rather to link a molecular description to another one — in terms of

²⁸ Kaplan, *op. cit.*

²⁹ A. Love, “Aspects of reductive Explanation in Biological Science: Intrinsicity, Fundamentality and Temporality”, *British Journal for the Philosophy of Science*, vol. 62, 2011, pp. 519-549.

structural and dynamic properties of networks — located at a different level of organization.”³⁰ Again, the variety of descriptions, and the levels at which the properties that these descriptions capture are assigned, become less problematic only once we accept the instrumentalist stance.

Potochnik tries to avoid the instrumentalist conclusion, while preserving the stronger, inference-based sense of explanations.³¹ For instance, one can invoke a notion of organism to explain a particular phenomenon (e.g., reproduction), but also one can treat it as a protein complex—the elements of which interact with the DNA sequences. The argument is that the lower-level explanations concern the lower-level properties — e.g. the explanations concerning molecular interactions between DNA and proteins — even when the target is the higher-level event, e.g. the development of a particular organismal trait. The traits of the phenotype are often explained by invoking both processes at both levels — e.g. molecular and environmental — constituting exactly what we call integrative explanations. This merging of levels supposedly does not constitute a difficulty, even though explanations are level-dependent, because competing higher-level and lower-level explanations *apply to different domains* and thus establish different types of generality. Thus, even though there is no supervening relation between phenotype and genotype, we can explain phenotype by processes at multiple levels.

It is not clear, however, what constitutes the explanatory power of such an integrative explanation and whether, if characterised this way, biological explanations are anything but a pile of mutually disconnected postulations concerning essentially disconnected processes. If accurate, this characterisation of biological explanations is the reason why philosophers flock to the instrumentalist camp in the first place, as it clearly exposes the explanations’ ontological superficiality.

Cartwright develops *pluralist realism* in another attempt to bypass the commitment to the reductionist principles without giving up on the inference-based sense of explanations.³² Its premise is that we can be realists without worrying about discrepancies between levels and domains. Yet this only seems to postpone the answer to the reductionist charge of incoherence and does not make instrumentalist standpoint less appealing.

³⁰ M. Morange, *History of Molecular Biology*, John Wiley and Sons, 2009, p. 23.

³¹ A. Potochnik, “Levels of Explanation Reconceived”, *Philosophy of Science*, vol. 77, 2010, pp. 59-72.

³² N. Cartwright, “Metaphysics of the Disunified World”, *PSA Proceedings*, 1994, pp. 357-364.

Dupré has argued for a metaphysical category of *process* as ontologically primary to individuals, thus leaving the aims and scope of both the theory and explanation to conveniently determine the boundaries of individuals.³³ It eliminates RE by eliminating entities as ontologically relevant. It thus effectively introduces the pragmatic choice of ontology in much the same spirit as instrumentalism.

Nevertheless, can process-based identity help us to come up with a satisfying account? Dupré argues that, based on everything we know about biological phenomena, it is likely that nature did not produce uniquely determined boundaries but only a flux of transient loci. He states that individuals should be treated as object-like abstractions from the flux of biological processes — i.e. temporarily stable nexuses in the flow of upward and downward causation. All components are “subtly different in each iteration but similar enough to constitute a distinctive process.” The processes, rather than the entities, are distinctive. Yet, *the real challenge is to characterise entities in the context of such processes: they are not simply loci, but loci and objects, as the cases we have invoked make fairly obvious*. The seemingly competing explanations, which focus on entities or processes, both contribute to the overall explanations, as in the case of plant biology. Or, they invoke both — as does the molecular mechanism of cis- and trans-regulation, where entities need to be defined in a precise way even though they are part of an overall regulative loop that prevents us from pronouncing any of them to be fundamental with respect to the other ones. Thus the question is to what extent we should adopt the main points of process metaphysics while avoiding oversimplification of the context of biological explanations.

4. Deriving central elements of a coherent ontology from multi-level integrative biological explanations

It is thus common to instrumentalism and other instrumentalist-minded alternatives that they do not question the approach that necessarily accepts metaphysical principles of reductionism as the foundation of ontological coherence. Rather, they dance around it in one way or another, while agreeing that the principles defining the coherence of ontology,

³³ J. Dupré, *Processes of Life: Essays in the Philosophy of Biology*, Oxford, Oxford University Press, 2012.

as defined by reductionism, are inadequate to explain the foundation of biological phenomena. As we have seen, biological explanations may be manipulation devices, but they often entail coherent and ontological frameworks of wide domains, providing that *the necessity of the reductionist criteria of coherence predicated on RE is not taken for granted*. Thus, a set of principles different from those established by reductionist metaphysics could yield a more adequate coherent ontology. The key difficulty, then, is to clarify the idea that organisms are entities that can be divided meaningfully in infinitely many ways, and to do it in such a way that it yields a coherent ontology rather than an instrumentalist-minded account. This task raises two questions. First, why should we try to achieve this, and second, is this achievable at all?

With respect to the first question, we should not overlook the fact that there is an empirical motivation for achieving such a goal. Introduction of causal reciprocity in biological explanations is empirically well-motivated across the domain of biological phenomena, even if it reflects a theoretical conundrum. In other words, *nature pushes us consistently in the direction of complex integrative explanations that defy conceptualization predicated on RE*.³⁴

With respect to the second question, the integrative biological explanations might not tell us much about the constitution of life; not necessarily because of an alleged ontological incoherence, but because of the pessimistic meta-induction problem. Our goal is more modest: it is to explore the resource of coherent (acceptable) ontological concepts of causal interactions, levels, and entities, as they are derived from biological explanations. *RE may be irrelevant to the explanatory power of integrative explanations to start with: our task is to find out the source of that power, if it is not merely success in manipulating organisms*. RE may or may not be part of the ontological structure of life. An adequate ontology may turn out to be reductionist, or perhaps more holistic, and postulate continuous non-atomic entities and their interactions. Yet, we will show that there is a coherent ontological aspect that is part of either, and that can be derived from integrative biological explanations.

Why should ontology be treated as anything more than an explicit explanatory framework in the first place? Why it should be treated as anything other than a heuristic device? Perhaps it is not more than a coherent heuristic framework — an ontology with

³⁴It is possible that the evolution of reproductive criteria of individuation we have mentioned above may evade these dilemmas, but the dilemmas concern an indispensable and central aspect of an overall knowledge of biological individuals and causes.

a small “o” — but our expectations, opposite to those of the instrumentalists, is that we can infer it from actual explanations, and that we expect it to be coherent.

One important reason why we should do our best to infer a coherent ontological framework from biological explanations has to do precisely with the aspect of biology instrumentalists focus on, namely its practice. *When biologists get clear on the ramifications of their ontology it can lead them to ask more pertinent questions, do it more efficiently and tackle phenomena in novel ways.* This is actually already a feature of biological practice, whether implicitly or explicitly. And even if it turns out that the foundations of biological theory are inconsistent, it is a useful and potentially productive ontological insight — simply assuming this point, or giving up on such a search too early, may be a missed opportunity.

As far as potential philosophical gains are concerned, we may be able to better understand what constitutes a coherent ontology in general, and how the domain of ontological research can be extended by dispelling false dichotomies introduced by reductionists. It may also be instrumental in realising the limits and the role of speculative, *a priori* search for fundamental distinctions and principles.

So what are biological entities, levels, and interactions across these levels, given that there are no fixed entity hierarchies? And what sort of entities can realise reciprocal interactions across levels in a coherent manner?

First, we need to come up with a minimal requirement for an entity that cannot be violated. It could be simply a spatio-temporally localized quantity of matter. Bio-entities are not uniquely defined but they are localised. Also, an entity must be causally *bound* internally.³⁵

Secondly, the key feature of biological entities that has been typically neglected is that they are *organized heterogeneous wholes*.³⁶ Even though this is a fairly obvious feature, it raises the question of how we distinguish them from homogenous wholes, and also puts into question the ways in which they can be analyzed and understood as genuine entities, given that homogenous entities we encounter in physics (e.g. elementary particles

³⁵Campbell and Bickhard offer an interesting account along similar lines, but it suffers from a problem that many other accounts suffer from: they attempt to establish an autonomy of macro-entities rather than focusing on their interrelations with micro-entities. Campbell R. and Bickhard M.H., “Physicalism, Emergence and Downward Causation”, *Axiomathes*, vol. 21, 2011, pp. 33-56.

³⁶Evan Thompson emphasizes the complexity of biological entities and their interactions, but does not raise the same questions we will raise. E. Thompson, *Mind in Life*, Cambridge Mass., MIT Press, 2007.

or rigid objects made of the same material) are presumed to be such entities. It is essential, given the nature of integrative biological explanations, both to establish basic natural entities, perhaps in a different sense, and to explore the ways in which they are related. Reductionist metaphysics completely disregards the relevance of this distinction.

Now, how does one establish a principle of individuation for organised wholes? One possibility is to leverage it via levels of heterogeneity. An appropriate measure of their individuality may be also a degree of coordination among parts.³⁷ Perhaps any criteria for individuating living organised wholes may be of a different kind than those commonly applied to homogenous entities.

This would further entail that criteria for individuating interactions be of a different kind as well. How do homogenous and heterogeneous entities mix? How do their parts interact and how do they interact as wholes? When a complex entity is established it can have causal homogeneity in acting on other entities analogous to homogenous entities. This implies that there are at least two sorts of interactions between homogenous and heterogeneous entities inherent to biological explanations. For example, is a piece of food entering metabolism a part of the metabolic system? What are the proper parts of metabolism, the single molecules or molecular complexes? Thus, when that piece of food interacts with other proteins should such interaction be treated as an interaction taking place among homogenous entities? But when we take into account that such interaction is mediated by the regulating mechanism of the system, then it is more adequate to treat it as an interaction between a homogenous entity and an organized whole. It is also not a trivial question whether protein-protein interactions should be characterised as interacting homogenous, or heterogeneous, organized wholes, and in what context. Similarly, genes and developmental processes are homogenous compared to the systems they interact with, while they are heterogeneous at the molecular level.

Because biological explanations, all the way to the molecular level, include kinds of interactions as basic theoretical points, the expectation that there is a basic explanatory level constituted of homogenous entities and their interactions seems implausible. *The trick is, rather, to integrate and coherently explicate the relationship between them, as*

³⁷ Not even physics exclusively studies homogenous wholes (even though it presupposes them as basic). See J. Wilson, "Ontological Butchery: Organism Concepts and Biological Generalizations", *Philosophy of Science*, vol. 67, 2000, p. 302.

well as their constitutive individuals. In achieving such a goal the existing notions of individuation and of entities have a very limited use. Wilson's (2000) suggestion of absolute and relative continuity may be helpful. Also, the relative extent of heterogeneity of an entity can be measured in terms of diversity of elements and their interactions.^{38, 39}

Given such desiderata, can we hold on to the concept of causation as a relation between entities realised across levels? Actually, the causal integration is the most prominent feature of organised wholes because it occurs *via* causal entanglements across levels. Thus, it is an essential part of biological explanations. Biological causal autonomy stems primarily from causal entanglements across reciprocally interacting levels (and entities of heterogeneity and organisation defined relative to the levels) rather than from homogenous determinate entities.

Thus, firstly, transient *entities* (as well as levels) are characterised by the extent of organisation and homogeneity. There are no fixed micro/macro entities, no basic micro-realiser of causal powers. And the entities are always contextualised by the criteria established by the following questions: 1) Is an entity analysed as a homogenous or a heterogeneous entity? 2) What kind of entity does it interact with?⁴⁰

Second, *causal relations (processes) with respect to the inter-level context, are treated only as tools of analysis.* In reality they are, typically, reciprocally interdependent — even though analysis can pick up dominance or an extent of relevance (not determination) of one causal direction.

Third, causal powers of individual fixed entities should be replaced by causal entanglements of levels and entities, characterised in terms of their *causal relevance*.

Working out such an account will be demanding, but given the possible benefits and the nature of available alternatives it will surely be worthwhile.

³⁸This was suggested to me by Stuart Kauffman.

³⁹The problem with defining individuals by functional integration is that they are usually integrated in still other integrated individuals and hierarchies — the properties that determine whether an individual is integrated hold to different degrees. It is hard to determine whether an individual is an organism using this criterion.

⁴⁰We will have two different types of mutually irreducible causes: *structural (dynamic)* and *efficient* (the former being something like Aristotle's formal cause devoid of essentialism).

References

- Bechtel, W. 2010, "The cell: locus or object of inquiry?", *Studies in History and Philosophy of Biological and Biomedical Sciences* 41, pp. 172-182.
- Brosius, J. 2009, "The Fragmented Gene", *Annals of the New York Academy of Sciences*, vol. 1178, 186-193.
- Campbell, R. and Bickhard, M.H., "Physicalism, Emergence and Downward Causation", *Axiomathes*, 21, pp. 33-56.
- Cartwright, N. 1994, "Metaphysics of the Disunified World", *PSA Proceedings*, pp. 357-364.
- Crane, J. 2012, "Biological-mereological Coincidence", *Philosophical Studies*, 161(2), pp. 309-325.
- Craver, C. 2007, "Constitutive Explanatory Relevance", *Journal of Philosophical Research* 32, pp. 3-20.
- Dawkins, R. 1989, *The Selfish Gene*, 2nd edition. Oxford: Oxford University Press. [First edition 1976]
- Griffiths, P.E. and Stotz, K. 2007, "Gene", In Hull, D.L. and Ruse, M. (eds.), *The Cambridge Companion to Philosophy of Biology*, Cambridge, Cambridge University Press.
- Dupre, J. 2012, *Processes of Life: Essays in the Philosophy of Biology*, Oxford University Press.
- Dupre, J. and O'Malley, A. O. 2007, "Metagenomics and biological ontology", *Studies in History and Philosophy of Biological and Biomedical Sciences*, 38, pp. 834-846.
- Forterre, P. 2010, "Giant Viruses: Conflicts in Revisiting the Virus Concept", *Intervirology*, 53: 362-378.
- Honig, B and Rohs, R. 2011, "Flipping Watson and Crick", *Nature* 470, pp. 472-473.
- Jacobson, K., Mouritsen, O. G. and Anderson R. G.W., "Lipid rafts: at a crossroad between cell biology and physics", *Nature Cell Biology*, vol. 9, n° 1, 7-14.
- Kagey, M.H. et al. 2010, "Mediator and cohesion connect gene expression and chromatin architecture", *Nature* 467, pp. 430-435.
- Kaplan, D. 1992, "The Relationship of Cells to Organisms in Plants: The Problem and Implications of Organismal Perspective", *International Journal of Plant Sciences* 153, pp. 28-37.
- Kauffman, S. A. 1970, "Articulation of Parts Explanation in Biology and the Rational Search for Them", *PSA Proceedings* 1970, pp. 257-272.
- Kim, J. 1993, *Supervenience and Mind*, Cambridge, Cambridge University Press.
1998, *Mind in a Physical World*, Cambridge, MIT Press.
2005, *Physicalism or Something Near Enough*, Princeton and Oxford, Princeton University Press.

- Lesne, A. and Victor, J.-M. 2006, "Chromatin fiber functional organization: Some plausible models", *The European Physical Journal E*, 19, pp. 279-290.
- Love, A. 2011, "Aspects of reductive Explanation in Biological Science: Intrinsicity, Fundamentality and Temporality", *British Journal for the Philosophy of Science* 62, pp. 519-549.
- Maynard Smith, J. 2000, "The Concept of Information in Biology", *Philosophy of Science* 67, pp. 177-194.
- Morange, M. 2009, *History of Molecular Biology*, John Wiley and Sons.
- Morange, M. 2009, "Articulating Different Modes of Explanation in Biology: The Present Boundary in Biological Research", Mapping the Future of Biology, *Boston Studies in Philosophy of Science* 266, pp. 15-26.
- Perovic, S. 2007, "The Limitations of Kim's reductive Physicalism in Accounting for Living Systems and an Alternative Nonreductionist Ontology", *Acta Biotheoretica* 55, pp. 243-267.
- Perovic, S. and Miquel P-A. 2011, "On Gene's Action and Reciprocal Causation", *Foundations of Science* 16, pp. 31-46.
- Potochnik, A. 2010, "Levels of Explanation Reconceived", *Philosophy of Science* 77, pp. 59-72.
- Rosenberg, A. and Kaplan, D. M. 2005, "How to Reconcile Physicalism and Antireductionism about Biology", *Philosophy of Science*, 72, pp. 43-68.
- Sitte, P. 1992, "A Modern Concept of the "Cell Theory": A perspective on competing hypotheses structure", *International Journal of Plant Sciences*, 153, S1-S6.
- Thompson, E. 2007, *Mind in Life*, MIT Press.
- Waters, K. 2008a, "How Practical Know-How Contextualizes Theoretical Knowledge: Exporting Causal Knowledge From Laboratory to Nature", *Philosophy of Science* 75, pp. 707-719.
- Waters, K. 2008b, "Beyond theoretical reduction and layer-cake anti-reduction." *Phil-sci Archive*. <http://philsci-archive.pitt.edu/3834/>.
- Wilson, J. 1999, *Biological Individuality: The Identity and Persistence of Living Entities*, New York: Cambridge University Press.
- Wilson, J. 2000, "Ontological Butchery: Organism Concepts and Biological Generalizations", *Philosophy of Science* 67: S301-S311.
- Wilson, D. S. and Elliott Sober, 1989, "Reviving the Superorganism", *Journal of Theoretical Biology* 136, pp. 337-356.
- Wimsatt, W. 2000, "Emergence as non-aggregativity and the biases of reductionism", *Foundations of Science* 3, pp. 269-297.

Épistémologie historique

L'épistémologie française et le problème de l'objectivité scientifique*

Les traits communs qui se trouvent dans l'œuvre épistémologique de G. Bachelard, G. Canguilhem et M. Foucault ont permis à J.-F. Braunstein d'identifier un certain " style français en philosophie des sciences. " ¹ Après avoir délimité certaines règles méthodologiques, adoptées par les trois auteurs pour rendre compte des discontinuités de la pensée scientifique et permettant d'identifier une allure spécifique de *l'épistémologie à la française*, je souhaite appliquer ces règles au problème de l'objectivité dans l'histoire des sciences. Mon objectif est de dégager " les conditions historiques de l'objectivité " et de montrer que le concept d'objectivité a connu, au cours de l'histoire de son usage, d'importantes transmutations qui méritent d'être prises en considération pour rendre compte de la polyvalence du mot dans ses multiples usages actuels. Cette excursion historique prend appui sur le monumental ouvrage de Lorraine Daston et Peter Galison intitulé *Objectivity*² dans lequel les deux historiens de la science montrent comment le concept d'objectivité scientifique s'est progressivement imposé, de façon polémique, contre d'autres vertus attribuées à la démarche scientifique, telles que la " vérité d'après nature ". En développant les conséquences épistémologiques de cette histoire de l'objectivité, je souhaite notamment défendre la thèse selon laquelle l'entreprise d'historiciser l'objectivité, loin de la rendre caduque ou de la relativiser au point d'en vider le contenu, vise au contraire à démontrer que les sciences ne peuvent produire des connaissances " objectives " qu'à condition d'établir leurs propres normes de productions des vérités et des faits. Je m'attacherai également à montrer comment l'histoire de l'objectivité nous permet de repenser la distinction traditionnelle entre *le sujet connaissant* et *l'objet connu*. Enfin, je vais formuler quelques critiques vis-à-vis de l'approche adoptée

* Article publié avec le soutien de la Fondation pour la Science de la République tchèque (*Czech Science Foundation*), GA14-07043S " Le Tournant pragmatiste en phénoménologie " .

¹ J.-F. Braunstein, " Bachelard, Canguilhem, Foucault. Le style 'français' en épistémologie ", in : P. Wagner (dir.), *Les philosophes et la science*, Paris, Gallimard, 2002, p. 920-963.

² L. Daston, P. Galison, *Objectivity*, Zone Books, New York 2007.

par Daston et Galison, qui semblent réduire la question de l'objectivité de la recherche scientifique à l'histoire d'une juste *représentation* de la nature.

Plutôt que de résoudre le problème de l'objectivité des résultats scientifiques dans le cadre d'une traditionnelle "théorie de la connaissance", les auteurs tels que G. Bachelard, G. Canguilhem et M. Foucault analysent les gestes concrets, les décisions et les pratiques par lesquels les différentes disciplines s'efforcent d'atteindre un savoir objectif. Les objets de la science ne préexistent pas aux différentes pratiques méthodiques qui les saisissent à travers des conceptualisations toujours nouvelles, en introduisant des discontinuités dans notre façon d'objectiver le monde. Autrement dit, la tâche de l'épistémologue est de reconnaître, à travers les démarches des savants, comment les problèmes et les objets scientifiques se sont constitués simultanément.

L'épistémologie de Gaston Bachelard se veut ainsi résolument non cartésienne : si l'on veut comprendre la nature de la rationalité, il faut s'intéresser à l'histoire des problèmes vis-à-vis desquels la raison a pu manifester sa puissance en proposant des solutions nouvelles. A l'encontre de Descartes et de Kant, Bachelard soutient que les structures de la raison ne se manifestent pas dans les principes abstraits qu'on pourrait dégager en forçant la raison à se retourner sur elle-même, mais dans les formes concrètes de son application. Or, c'est la science qui représente la forme par excellence des applications successives de la pensée à des problèmes concrets. Par conséquent, il faut réfuter les conceptions purement philosophiques qui établissent une forme idéale de la rationalité, car celles-ci non seulement se renvoient dos-à-dos dans leurs divergences respectives, mais surtout elles ne rendent justice ni au développement incessant des sciences, ni aux ruptures épistémologiques auxquelles ce développement polémique donne lieu. La réflexion philosophique sur la science doit reposer sur l'analyse détaillée du processus discontinu, à travers lequel tel ou tel domaine du savoir s'efforce d'atteindre le statut de science objective.

Or, cette conquête est conçue par Bachelard comme une lutte : chaque science est toujours dans un rapport polémique avec elle-même, elle s'établit en tant que science en luttant contre son passé le plus récent et contre les obstacles et les erreurs qui hantent cet

héritage. Ni les observations des faits, ni les protocoles expérimentaux ne se limitent pas à une pure constatation de l'objectivité qui serait déjà là : “ L'observation scientifique est toujours une observation polémique ; elle confirme ou infirme une thèse antérieure, un schéma préalable, un plan d'observation. ”³ Et si l'on passe de l'observation à l'expérimentation, ce rapport conflictuel vis-à-vis des connaissances préalables ne fait que s'exacerber : l'expérimentation non seulement filtre et trie les phénomènes déjà connus, mais également produit des phénomènes jamais constatés jusque-là moyennant des techniques nouvelles et des instruments construits pour l'occasion.⁴ Enfin, lorsqu'un nouveau concept ou principe explicatif se met en place, les anciens concepts apparaissent comme des obstacles à surmonter, comme des idées fixes dont il faut s'émanciper. L'exemple le plus connu est celui du “ phlogistique ” qui constituait – jusqu'à l'éclaircissement du rôle de l'oxygène dans la combustion par Lavoisier à la veille de la Révolution – un obstacle épistémologique à une théorie rationnelle de la combustion.

Cependant, ce ne sont pas seulement les théories scientifiques erronées (formées autour de concepts désormais périmés tels que le phlogistique, le calorique ou l'éther), mais aussi des méthodes ayant auparavant contribué de façon significative à l'avancée de la science qui peuvent se figer dans de nouveaux obstacles épistémologiques. Ainsi, la mise en valeur de l'observation directe, qui a favorisé la rupture avec la physique aristotélicienne au 17^e siècle, est devenue à son tour l'obstacle pour l'avancée de la théorie atomiste de la matière. De même, la philosophie de la science peut être source d'obstacles épistémologiques, dans la mesure où elle a tendance à canoniser les aspects contingents d'une certaine époque, comme s'il s'agissait des structures immuables du monde et de la raison. Selon Bachelard, ce penchant qui nous porte à tenir pour nécessaire ce qui n'est que contingent se manifeste dans la *Critique de la raison pure* où Kant a absolutisé la science de Newton, ses concepts et les règles d'usage de ces concepts, comme s'il s'agissait du canon même de la rationalité scientifique.⁵ C'est pour cette raison que Bachelard rejette résolument toute *théorie de connaissance* au sens kantien du terme : les “ conditions de possibilité ” invariantes ou universelles de la connaissance en général sont une *chimère* des philosophes dont nous devons nous défaire si nous souhaitons saisir le mouvement des ruptures, à travers lesquelles les différentes théories scientifiques ont recours à différentes formes de rationalisation et de formalisation de leurs objets d'études.

³, Gaston Bachelard, *Le nouvel esprit scientifique*, Paris, P.U.F., 1963, p. 12.

⁴ *Ibid.*

⁵ cf. Gaston Bachelard, *La philosophie du non*, Paris, P.U.F., 1940, p. 30.

Il s'en suit que la rationalité scientifique se conjugue au pluriel, qu'il n'y a pas de règles générales qui pourraient délimiter la scientificité, tout comme il n'y a pas de normes qui pourraient départager une fois pour toutes la rationalité de l'irrationalité. Les normes de scientificité s'instituent et se transforment dans le cadre des disciplines singulières et il revient à l'épistémologue de déceler cette normativité dans l'analyse de tel ou tel domaine du savoir dans son développement historique discontinu. Par conséquent, il n'existe pas de méthode universellement applicable à tous les objets de la recherche scientifique, mais seulement des méthodes multiples, comme le souligne Canguilhem, en prolongeant la critique bachelardienne des recherches visant à postuler un modèle universel de la formation du savoir :

“L'épistémologie de Gaston Bachelard [...] repousse l'idée commune, à quelques nuances près, à A. Comte, à Chevreul, à Cl. Bernard, selon laquelle il existe une méthode positive constituée des principes généraux dont seule l'application est diversifiée par la nature des problèmes à résoudre.”⁶

L'épistémologie française prend ainsi en considération non seulement le conditionnement historique de la rationalité scientifique, mais aussi la pluralité des rationalités qui peuvent s'imposer en parallèle ou se juxtaposer sur le plan synchronique. C'est Foucault qui a su résumer mieux que tout autre ces deux idées essentielles de Bachelard, en posant l'historicité radicale de la rationalité scientifique comme l'objet d'une histoire critique des sciences et la pluralité des rationalités comme l'objet d'une géographie des sciences. Dans son introduction à l'édition américaine du *Normal et le pathologique* de Canguilhem, Foucault écrit :

“On a posé à la pensée rationnelle la question non seulement de sa nature, de son fondement, de ses pouvoirs et de ses droits, mais celle de son histoire et de sa géographie, celle de son passé immédiat et de ses conditions d'exercice, celle de son moment, de son lieu et de son actualité.”⁷

Alors que l' “ histoire de la pensée rationnelle ” vise à dégager la constitution d'une discipline ainsi que les événements, les gestes et les partages qui ont rendu possible notre savoir présent et les discontinuités qui remettent en cause l'idée d'une cumulation des connaissances, la tâche d'une “ géographie de la pensée rationnelle ” est plus difficile à

⁶ George Canguilhem, *Études d'histoire et de philosophie des sciences*, Paris, Vrin, 1994, p.166-167.

⁷ Michel Foucault, „Introduction“, in: Georges Canguilhem, *On the normal and the pathological*, Dordrecht, Boston, & London, D. Reidel, 1978, p. ix-xxi; repris dans Michel Foucault, „La vie : l'expérience et la science“, *Revue de métaphysique et de morale*, vol. 90, n° 1, 1985, p. 3-14, p. 6.

cerner. Afin de préciser le sens accordé par Foucault à cette “ géographie ”, nous pouvons délimiter l’objet d’une telle entreprise par les questions suivantes : Quelle était la configuration spatiale (des États, des universités, des cliniques ou des laboratoires) qui a permis de constituer les divers champs de savoir dans leur autonomie relative ? Quelles démarcations des frontières fallait-il établir pour que la médecine, la sociologie, la démographie ou la psychologie puissent prétendre au statut de sciences ? Quels procédés fallait-il introduire pour que la vie, la pensée et le comportement des individus, mais aussi des populations puissent devenir l’objet d’une connaissance objective ? A partir de telles questions, on voit bien que l’objectif de l’épistémologie historique consiste à atteindre *les conditions qui rendent possible la science*, c’est-à-dire un objectif semblable à celui du questionnement poursuivi par Kant dans la *Critique de la raison pure* : Quelles sont les catégories et les contraintes qui, sans découler directement de la recherche empirique, rendent possibles des énoncés empiriques, en permettant de penser telle ou telle idée ? Cependant, à l’encontre de Kant, ces conditions sont considérées comme historiquement données et dépendant du contexte local et contingent de leur émergence. Il en découle que l’histoire de la science n’est pas constituée seulement des rectifications d’erreurs et d’accumulations et de transformations des connaissances, mais qu’à travers les changements radicaux de méthodes et de conceptualisations, nous sommes confrontés à des transformations incessantes de notre compréhension *de ce qui peut ou ne peut pas être* considéré comme un savoir.

“ Le travail de Koyré, de Bachelard, de Cavaillès et de Canguilhem [pose des interrogations] à une rationalité qui prétend à l’universel tout en se développant dans la contingence ; qui affirme son unité et qui ne procède pourtant que par modifications partielles.”⁸

Ainsi, ce qui apparaissait comme une contrainte *a priori* de la pensée se trouve, notamment chez Foucault, ramené à des conditions contingentes liées aux vicissitudes du développement d’une discipline.

Or, si les conditions *a priori* de la science ne découlent aucunement des structures anhistoriques de la raison, comme nous l’avons montré ci-dessus, le sujet de la science ne saurait être identifié à une conscience individuelle, capable de juger par elle-même de l’objectivité de ses propres contenus, mais il faut le chercher dans le sujet collectif et historiquement situé de la communauté des chercheurs partageant les mêmes normes.

⁸ Foucault, “ La vie: l’expérience et la science ”, p. 6.

Selon Bachelard, les normes qui doivent assurer l'objectivité de la connaissance scientifique, ne peuvent aucunement être établies par un *cogito* individuel, mais seulement par le *cogitamus* pluriel de la " cité scientifique ".⁹ La communauté de ceux que Bachelard nommait " les travailleurs de la preuve " devient alors l'instance qui répond de la validité des normes de la connaissance scientifique, et par là même de l'objectivité des résultats atteints en suivant de telles normes. Sur ce point, il serait possible d'objecter à Bachelard d'avoir reproduit, à son insu, le présupposé selon lequel les normes de l'objectivité résultent de l'activité de la conscience, même si celle-ci est envisagée désormais dans sa dimension collective. Il est possible d'entrevoir cette tendance dommageable à retomber dans le psychologisme notamment dans *Le rationalisme appliqué* où Bachelard situe la normativité de la recherche scientifique dans le " surmoi intellectuel " auquel appartient d'effectuer une " surveillance intellectuelle de soi ".¹⁰ Une telle surveillance repose sur le dédoublement du " moi " du chercheur, qui poursuit sa recherche en se penchant sur soi-même, afin de mieux contrôler la justesse des résultats, le respect des méthodes reconnues par la cité scientifique, de même que les choix méthodologiques compte tenu du problème à résoudre (ce qui peut éventuellement déboucher sur une révision des règles du jeu). Canguilhem considère une telle description de l'origine des normes et de leur révision comme ingénieuse, mais insuffisante :

"Fonder l'objectivité de la connaissance rationnelle sur l'union des travailleurs de la preuve, la validité du rationalisme sur la cohésion d'un corrationnalisme ; fonder la fécondité de *mon* savoir sur la division du *moi* en moi d'existence et moi de surexistence, c'est-à-dire au sein d'un *cogitamus*, toute cette tentative est ingénieuse, convaincue, mais non pleinement efficace à convaincre."¹¹

En effet, en établissant les normes de la connaissance scientifique dans le partage des raisons individuelles au sein d'une communauté, n'est-on pas amené à les réduire aux facteurs d'ordre psychologique ? Même si nous sommes enclins à souscrire à la thèse de Bachelard, selon laquelle l'objectivité dépend du consensus de la cité scientifique, il est primordial de préciser le sens d'une telle dépendance. Si nous voulons comprendre l'effort de la communauté scientifique pour atteindre des connaissances qui ne seraient pas dépendantes de la perspective subjective de tel ou tel chercheur, il faut avant tout prendre en considération les conditions historiques et institutionnelles qui précèdent les

⁹ Gaston Bachelard, *Le rationalisme appliqué*, Paris, P.U.F., 1966, p. 57.

¹⁰ *Idem*, p. 75.

¹¹ Georges Canguilhem, *Études d'histoire et de philosophie des sciences*, Paris, Vrin, 1983, p. 205.

performances cognitives et expérimentales des scientifiques s'efforçant d'élaborer une version aussi épurée que possible du savoir objectif.

C'est Foucault qui pose ces conditions historiques du savoir comme thème majeur de ses recherches multiples dans le domaine de l'archéologie du savoir. Sa notion de l' "*a priori* historique " semble en effet ouvrir une voie prometteuse pour traiter du problème de l'objectivité et de ses avatars dans l'histoire des sciences. Foucault introduit la notion de l' "*a priori* historique " dans sa *Préface aux Mots et les choses* :

“ [...] une étude qui s'efforce de retrouver à partir de quoi connaissances et théories ont été possibles ; selon quel espace d'ordre s'est constitué le savoir ; sur fond de quel *a priori* historique et dans l'élément de quelle positivité des idées ont pu apparaître, des sciences se constituer, des expériences se réfléchir dans des philosophies, des rationalités se former, pour, peut-être, se dénouer et s'évanouir bientôt. ”¹²

Dans le contexte des recherches archéologiques poursuivies par Foucault dans les années 1960, nous pouvons comprendre de tels *a priori historiques* comme des matrices inconscientes gouvernant l'espace des énoncés possibles, qui apparaissent dans les écrits scientifiques de telle ou telle période historique. Ils sont conçus comme les conditions de possibilité de savoir d'une culture scientifique donnée, comme un ensemble de règles rendant possibles des énoncés de la science. Mais alors que chez Kant, les notions d' "*a priori* " et d' "*historique* " s'opposent, Foucault ne recule pas devant un tel oxymoron.¹³ L'archéologie du savoir se propose ainsi de dégager les conditions dans lesquelles un énoncé peut prétendre à son objectivité, tout en analysant les principes selon lesquels de tels énoncés subsistent, se transforment et finalement disparaissent de l'horizon de ce qui peut être objectivement déclaré.

Malheureusement, Foucault ne parle presque jamais des *a priori* de l'objectivité dans les sciences naturelles, puisqu'il s'attache plutôt à analyser, de façon critique, la prétention des sciences dites humaines à constituer l'homme en tant qu'objet d'une connaissance scientifique. Cependant, les méthodes de l'épistémologie française en général – et de l'archéologie foucauldienne en particulier – ont servi de source d'inspiration à Lorraine Daston et Peter Galison qui ont consacré à l'étude des transformations du concept d'objectivité, au cours de plus de trois siècles de science occidentale, un volume

¹² Michel Foucault., *Les mots et les choses*, Paris, Gallimard, 1966, p. 13.

¹³ Comme le remarque l'auteur de l'*Archéologie du savoir* lui-même, “ juxtaposés, ces deux mots font un effet un peu criant ”, *Archéologie du savoir*, Paris, Gallimard, 1969, p. 167.

monumental, intitulé brièvement *Objectivity*. Les deux historiens de la science y dessinent les multiples péripéties, à travers lesquelles le concept de l'objectivité s'est finalement établi par opposition à celui de la subjectivité. Leur tableau panoramique restitue les différents idéaux qui ont réglé la volonté des savants d'aboutir à une saisie aussi fidèle que possible de la réalité et fournit suffisamment de matériau historique pour mieux cerner les *a priori historiques* qui ont permis l'émergence de la connaissance objective.

En abordant l'objectivité comme une notion historiquement variable, Daston et Galison poursuivent une vaste enquête sur la genèse et les mutations de ce concept, à travers une recherche sur les différentes façons de représenter les phénomènes naturels dans les atlas scientifiques. Considérons d'abord la thèse que les deux auteurs veulent réfuter :

“ Qu'elle soit comprise comme le point de vue de nulle part ou bien comme l'application d'une règle algorithmique, qu'elle soit louée comme le noyau même de l'éthique scientifique ou incriminée comme un détachement sans âme, l'objectivité est censée être abstraite, intemporelle et monolithique. ”¹⁴

Or, Daston et Galison nous montrent que l'objectivité n'est ni intemporelle, ni monolithique, mais qu'elle a sa propre histoire, scandée par un certain nombre de reconfigurations des catégories fondamentales de la rationalité scientifique. La reconstruction de cette histoire nous permet de mieux saisir la polyvalence actuelle du concept. En effet, nous désignons comme “ objective ” une connaissance qui s'en tient rigoureusement aux faits et qui n'est pas biaisée par nos préconceptions : dans ce sens, l'objectif serait associé à l'empirique. Cependant, le terme d'objectif est employé également comme synonyme du rationnel, au sens d'une évidence que tout être rationnel doit reconnaître. Outre cela, la notion de l'objectivité renvoie à une approche impartiale (c'est dans cet esprit que nous parlons d'un jugement objectif) aussi bien qu'à la capacité de prendre distance vis-à-vis de ses propres émotions. Enfin, nous employons ce terme également comme synonyme de la réalité en soi ou de l'existence des choses telles

¹⁴ Daston, Galison, *Objectivity*, p. 51. Nous traduisons.

qu'elles sont, indépendamment de toute projection de notre esprit sur eux.¹⁵ De toute évidence, le terme d'objectivité recèle un certain nombre de couches sémantiques qui ne peuvent pas être corrélées sans difficulté : ainsi, il n'est pas très facile de déterminer le dénominateur commun entre la capacité de maîtriser ses émotions et le présupposé d'un ordre neutre et fondamental de toute réalité. Si donc nous voulons comprendre les racines historiques de cette exubérance sémantique de l'objectivité et si nous voulons saisir l'enchevêtrement de toutes ces nuances, il faudra rechercher, dans l'histoire des sciences, le rôle assumé par le postulat du principe d'objectivité, tout en comparant et délimitant ce principe par rapport aux autres candidats pour une connaissance fidèle, adéquate et légitime des choses.

Selon la thèse fondamentale de Daston et Galison, les différentes époques de la science moderne se distinguent par une certaine “ vertu épistémique ” qui représente, pour les savants de telle ou telle période, l'ensemble des normes régissant la manière de connaître et de représenter la nature. Au XVIII^e siècle, l'idéal à atteindre, communément partagé par les savants, consistait dans une “ vérité d'après nature ”, l'objectif étant d'aboutir à une représentation aussi fidèle que possible de la nature en trouvant l'exemplaire idéalisé et caractéristique. Le principe de cette vertu épistémique impose au savant, accompagné de son illustrateur, de faire plutôt le portrait de l'espèce que d'un exemplaire particulier. Il s'agit pour lui de déterminer, à travers une représentation idéalisée, ce qui est caractéristique, universel et typique.¹⁶ L'attention du savant, guidée par la raison, permet de dompter la variabilité présente dans les exemplaires particuliers et de reconnaître, au-delà de l'imperfection des individus, les “ véritables ” espèces des plantes et d'autres organismes. Il s'agit ainsi d'une image synthétisée de la diversité empirique et non pas d'une représentation de tel ou tel individu auquel le savant était lui-même confronté durant ses recherches.

Cette “ vérité d'après nature ” a été progressivement remplacée par un autre idéal, lié à d'autres conceptions du travail à effectuer sur soi pour atteindre la vérité des choses. A partir du milieu du XIX^e siècle, la tâche de la science consiste à éradiquer toute projection subjective sur l'objet de la science. Cette vertu épistémique conduit dans la pratique à utiliser autant que possible des instruments d'enregistrement tels que la photographie, le daguerréotype et autres outils capables de représenter la réalité sans

¹⁵ Voir Daston, L., « Objectivity and the Escape from Perspective », *Social Studies of Science* 22, n° 4, 1992, p. 598.

¹⁶ Daston, L., Galison, P., *Objectivity*, p. 20.

l'interférence de l'homme. Il s'agit désormais d'immuniser l'objet étudié contre toute contamination qui pourrait venir de l'homme : comme le démontrent Daston et Galison sur de nombreux exemples de photographies de flocons de neige ou de cristaux, la mise en œuvre des procédés mécaniques de reproduction a pour but d'éliminer l'intervention humaine, d'immuniser l'objet étudié contre toute contamination qui pourrait venir de l'homme, c'est-à-dire de contourner le danger imminent qui consiste dans la projection de nos propres attentes, hypothèses et théories sur la réalité en déformant ainsi son image. C'est pour cette raison que les flocons de neige, les minéraux ou les individus vivants, à la différence des images idéalisées représentées dans les atlas du XVIII^e siècle, sont représentés avec toutes leurs particularités et asymétries : si les auteurs d'atlas de la seconde moitié du XIX^e siècle n'osaient pas effacer le moindre détail de l'image, leur abstention d'intervenir se fait au nom de l' " objectivité mécanique ".

Toutefois, les insuffisances et les apories liées à la constitution d'une objectivité toute mécanique émergent au tournant du XX^e siècle : non seulement il est impossible d'exclure entièrement la perspective du savant qui oriente la lentille de l'appareil, mais surtout, la communauté scientifique se rend compte de la nécessité de réintroduire les procédures de jugement pour donner une interprétation sensée des données brutes et pour corriger les limites et les insuffisances du procédé de pur enregistrement. L'objectivité mécanique laisse place, notamment à partir des années 1920, à ce que Daston et Galison appellent le " jugement entraîné ", permettant de dépasser les impératifs aporétiques de l'objectivité mécanique. Si les auteurs d'atlas du XX^e siècle se permettent de mettre en relief certaines structures ou formes significatives [*patterns*] selon leur jugement expert, ils le font au nom d'une conviction partagée, selon laquelle l'interprétation adéquate des images obtenues grâce aux télescopes, aux microscopes ou les enregistrements des électroencéphalogrammes et IRM, requiert un groupe d'experts attestés qui jouissent d'une crédibilité suffisante au sein de leur discipline. Ainsi, donner une image objective d'une lésion cérébrale ou d'un champ magnétique du soleil suppose de passer par un entraînement permettant d'exercer son regard à reconnaître les structures significatives, considérées comme dignes d'intérêt dans tel ou tel domaine de l'expertise.

En ce qui concerne le problème épistémologique de l'objectivité dans la science, il est possible d'extraire trois thèses fondamentales qui découlent du tableau panoramique dessiné par Daston et Galison. D'abord, l'histoire de l'objectivité permet de mieux cerner le rapport étroit qui relie les différentes conceptions historiques de “ la réalité telle qu'elle est ” avec les ruptures successives dans notre façon de concevoir le “ moi ”. Ensuite, l'excursus historique dans les différents régimes de la représentation et de la rationalisation de la nature permet de dégager, pour chaque période considérée, un éthos spécifique déterminant la manière dont il faut étudier, cerner et décrire les phénomènes naturels. Enfin, il est possible de montrer que l'entreprise visant à historiciser la notion de l'objectivité ne revient pas à rejeter la rationalité ou le réalisme des résultats scientifiques, mais à mieux délimiter la prétention de la science à expliquer les phénomènes en respectant les articulations dessinées par la nature elle-même.

Considérons d'abord la première thèse selon laquelle les différentes manières de rechercher une image fidèle, objective, ou correspondant à une expertise, de la nature ont émergé parallèlement aux apparitions de nouveaux modes de la subjectivité. L'ouvrage de Daston et Galison permet d'en préciser le contenu en montrant comment chaque forme de l'objectivité (ou de ses alternatives) est déterminée par la censure de l'un des aspects de notre subjectivité, considéré comme particulièrement dangereux à telle ou telle époque. L'objectivité, affirment Daston et Galison, “ implique la suppression d'un aspect du moi, et s'oppose à la subjectivité ”.¹⁷ Ainsi, au XVIII^e siècle, alors que le terme d'objectivité n'a pas encore reçu son sens actuel, il fallait supprimer ou dompter notamment son imagination effrénée. L'antidote recommandé consiste alors dans un usage raisonné de son esprit, qu'il est possible d'entraîner, à l'instar de nos muscles, par une expérience répétée, afin d'acquérir par là la capacité de saisir l'uniformité fondamentale de la nature derrière la diversité apparente de ses manifestations. La raison doit ainsi remettre l'imagination à sa place, résister à ses tentations et compenser la passivité des sens par une intuition raisonnée de la structure profonde et durable de la réalité.

Une conception nouvelle du moi actif, faisant suite à la révolution copernicienne de Kant, a introduit une nouvelle source d'angoisse dans la recherche scientifique: si “ les

¹⁷ *Ibid.*, p. 36.

objets se règlent sur notre connaissance ”¹⁸, comment empêcher à nos spéculations et interprétations de compromettre la validité des résultats de notre enquête ? La crainte majeure concerne alors la possibilité de déformer l’objet à connaître en projetant sur lui nos anticipations, nos hypothèses, nos perspectives. Certes, une telle crainte n’est aucunement basée sur une interprétation fidèle à l’esprit de l’argument kantien, mais elle résulte plutôt d’une réappropriation de Kant par les savants du XIX^e siècle. Ces usages et ces relectures de Kant donnent alors naissance à l’idée d’une science impersonnelle, une science dont la condition de possibilité consiste dans une abnégation du chercheur, dans une disciplinarisation de son propre “ moi ” afin de l’empêcher de projeter sur l’objet à connaître une partie de sa spécificité. Ainsi, selon Rudolf Virchow, promoteur d’une médecine rigoureusement orientée vers les sciences naturelles, la tâche du savant consiste à “ désobjectiver ” son moi, qui représente une menace pour la théorie scientifique.¹⁹ Le paradoxe d’une telle entreprise est pourtant manifeste : cette abnégation ne saurait être atteinte qu’au prix d’un effort héroïque de la volonté, c’est-à-dire par un acte souverain du sujet connaissant. Par conséquent, l’impartialité du savoir n’est pas seulement le fruit d’une connaissance de l’objet, mais d’un acte moral du sujet.

Cette considération nous amène à la seconde des thèses mentionnées plus haut, selon laquelle il est impossible d’introduire une coupure nette entre la forme des théories scientifiques et les principes éthiques qui leur ont permis d’exister, de même qu’il est impossible de concevoir la validité des résultats indépendamment des méthodes mises en œuvre pour les obtenir. Chaque régime de production scientifique des faits est ainsi animé par un type spécifique de principes normatifs qui déterminent non seulement le mode de représentation des images, mais aussi la nature de travail qu’il faut effectuer sur soi pour pouvoir accéder à la vérité des choses. Alors que les principes éthiques des naturalistes du XVIII^e siècle, insistant sur un entraînement prolongé de l’attention et de la mémoire, peuvent s’apparenter à un développement original de l’éthique aristotélicienne et néo-stoïcienne des vertus, l’éthique de l’objectivité du XIX^e siècle est dans une large mesure inspirée par celle de Kant, car la volonté est la faculté souveraine dont le savant doit faire un bon usage, s’il veut dompter son moi exubérant en lui empêchant de projeter ses attentes sur les objets de son étude. Il y a ainsi une forme d’ascèse chez le scientifique du XIX^e siècle, inséparable de sa méthode de travail. A la suite de Bachelard, de Canguilhem

¹⁸ Emmanuel Kant, *Critique de la raison pure*. Préface de la seconde édition, Paris, J. Gibert, 1946, p. 21.

¹⁹ Voir le discours de R. Virchow à l’occasion de *Versammlung deutscher Naturforscher und Ärzte* en 1871, cité par Daston et Galison, *Objectivity*, p. 189.

et de Foucault, Daston et Galison remettent ainsi en question l'un des présupposés fondamentaux de la philosophie des sciences, selon lequel l'objectivité est l'affaire d'une "théorie de connaissance", assimilable à l'impartialité et à la neutralisation de toute valeur. Bien plutôt, l'objectivité est un phénomène normatif qui connaît des transmutations historiques en fonction des différentes visées régulatrices, qui se succèdent dans les étapes du développement de la science moderne.

Nous pouvons toutefois nous demander si une telle moralisation et historicisation du terme d'objectivité ne conduit pas à un relativisme de principe. En effet, en délimitant l'objectivité comme une modalité historiquement et moralement conditionnée du rapport de la théorie à la nature, on refuse de lui attribuer l'universalité qui semble représenter le présupposé nécessaire de toute approche dite réaliste de la science. Quelles sont alors les conséquences proprement épistémologiques de la thèse selon laquelle l'objectivité n'a pas toujours été un principe régulateur de la recherche scientifique, puisqu'elle s'est développée seulement au milieu du XIX^e siècle et en corrélation avec un idéal assez paradoxal de l'effacement de soi ? Si, de plus, les deux historiens insistent pour dire que l'objectivité – en tant qu'idée régulatrice guidant la recherche dans les sciences – est intrinsèquement liée à une conception historique de la subjectivité, dérivée de la philosophie kantienne et de ses avatars, et sévèrement attaquée tout au long du XX^e siècle, on peut se demander si leur enquête ne revient pas, en fin de compte, à effacer l'objectivité, qui pourrait finalement disparaître comme "à la limite de la mer un visage de sable".²⁰ En d'autres mots, l'insistance sur les facteurs contingents et sur l'historicité du concept de l'objectivité n'amène-t-elle pas à considérer l'objectivité comme un artifice, comme une construction idéologique, comme une illusion ? Certes, les deux auteurs soulignent le caractère contingent de l'idéal d'objectivité ; toutefois, historiciser l'objectivité ne revient pas à la nier ou à la dénoncer comme une chimère du passé qui n'a plus droit de cité dans le travail de la science. Les idées normatives partagées, même lorsqu'elles s'attachent à un idéal irréalisable dans sa forme pure, amènent à des résultats concrets au niveau de la représentation de la nature et de la création de nouveaux objets de la science. Par ailleurs, Daston et Galison refusent de considérer l'approche constructiviste et l'approche réaliste de l'objectivité comme nécessairement contradictoires :

²⁰ Foucault, *Les mots et les choses*, p. 398.

“ Objecter que l’objectivité mécanique [...] est un leurre ou une illusion, car elle n’est jamais réalisée dans sa forme la plus pure, serait comme affirmer la même chose à propos de l’égalité ou de la solidarité. [Dans le cas du jugement exercé et de l’objectivité mécanique], il ne s’agit pas d’opposer l’idéologie à la réalité, mais de deux visions régulatrices de la science qui sont distinctes et parfois rivales et qui sont aussi réelles que les images qu’elles produisent, tout en étant elles-mêmes produites par des circonstances historiques spécifiques. ”²¹

Historiciser la notion d’objectivité ne revient donc aucunement à nier qu’elle existe, mais à s’interroger sur les catégories profondes de la rationalité scientifique, sur leur interdépendance et leurs reconfigurations au sein de l’histoire des sciences qui se distingue par une incessante créativité des normes. Daston et Galison renouent ainsi, de manière implicite, avec la thèse de Canguilhem, selon laquelle c’est précisément l’historicité qui représente le signe distinctif de la scientificité :

“ Une science qui n'a pas d'histoire, c'est-à-dire une science dans laquelle il n'y a pas récusation de certaines conditions d'objectivité à un moment donné et substitution de conditions d'objectivité plus objectivement définies, une discipline ainsi conçue n'est pas une science. ”²²

Il s’en suit que les conditions de possibilité des énoncés scientifiques doivent nécessairement être historiques, sans quoi la science serait réduite à une série infinie de discours mutuellement entrelacés et déduits à partir de prémisses invariables. L’historicité est ainsi ce qui distingue les sciences véritables des pseudo-sciences, telles que l’astrologie, dont le propre est de ne pas connaître d’histoire au sens bachelardien des ruptures épistémologiques.

En conclusion, je me permettrai d’exprimer quelques réserves sur la tentative de Daston et Galison pour rendre compte de l’objectivité par une histoire des transformations ayant lieu dans la production des images des phénomènes naturels. Une lecture attentive de cet ouvrage monumental donne progressivement lieu au soupçon qu’il s’agit davantage

²¹ Daston, Galison, *Objectivity*, p. 378. Nous traduisons.

²² Georges Canguilhem, “ Objectivité et historicité ”, in: J.-M. Auzias, *Structuralisme et marxisme*, Paris, Union générale d’éditions, 1970, p. 235.

d'une histoire de la *représentation* de la nature que d'une histoire de l'objectivité de la *connaissance* scientifique. Certes, les deux auteurs ont su démontrer l'importance des atlas dans l'orientation de notre regard et dans la représentation des phénomènes naturels. Et pourtant, le problème de l'objectivité de la connaissance scientifique ne se réduit pas à celui d'une juste représentation des cristaux, des plantes, des embryons ou des champs magnétiques. Ni les lois de la mécanique newtonienne, ni la théorie de la relativité d'Einstein ne sont de simples représentations de la nature. Si la notion de l'objectivité doit recevoir un sens non trivial, il est nécessaire de reconnaître qu'elle concerne moins une image fidèle des phénomènes, qu'une explication permettant de maîtriser la diversité des expériences par les principes unificateurs de la pensée. La généalogie de la rationalité scientifique consisterait alors dans la découverte des concepts, des catégories et des règles qui se sont imposés au cours de l'histoire des sciences européennes et qui ont permis de trier, de penser et d'unir les données de l'expérience dans un corpus de savoir objectif.

A l'encontre de la tentative pour expliquer la constitution du savoir objectif exclusivement dans les termes des “ gestes, techniques, habitudes et dispositions qui sont inculqués par l'entraînement et la répétition quotidienne, ”²³ il est possible d'objecter que l'étude d'une objectification progressive de la nature doit prendre en considération avant tout les différents aspects de la formalisation et de l'idéalisation des phénomènes par les sciences de la nature. Il serait ainsi souhaitable de corréler les contingences historiques, découvertes par l'épistémologie historique, aux thèses sur l'“ idéalisation mathématique ”, présentées par Husserl dans la *Crise des sciences européenne*. C'est que l'objectivité est inconcevable sans tenir compte des opérations telles que a) la procédure de mesure qui permet d'accéder à la dimension d'un sens identiquement valable pour tous ceux qui l'effectuent ; b) l'explication des phénomènes naturels par le principe d'une causalité pure, ayant pour conséquence une prédiction exacte des phénomènes futurs ; c) l'idée d'une mathématisation indirecte, qui présuppose de pouvoir trouver, pour chaque donnée qualitative des objets naturels (telle que la couleur, la chaleur ou le mouvement), un équivalent dans le domaine rigoureusement mathématique des formes spatio-temporelles (telles que la longueur d'onde, l'énergie cinétique et l'espace parcouru au cours d'un intervalle de temps).²⁴

²³ Daston, Galison, *Objectivity*, p. 52.

²⁴ Voir Edmund Husserl, *La crise des sciences européennes et la phénoménologie transcendantale*, Paris, Gallimard, 1976, II, § 9 c, p. 43 : “ Tout ce qui se révèle comme réel dans les qualités sensibles spécifiques

En faisant l’impasse sur ces opérations effectuées par la science moderne depuis son point de départ chez Galilée et en se limitant à l’histoire des représentations, Daston et Galison ne rendent pas entièrement justice à la prétention de la science à connaître la réalité dans son indépendance par rapport à celui qui la pense. Cette prétention se trouve finalement réduite à la forme contingente d’un éthos, historiquement délimité, d’une connaissance désintéressée et désobjectivée. Lorsque les deux historiens américains conçoivent l’objectivité comme le régime de production scientifique “ qui aspire à l’effacement de toute trace de celui qui est à l’origine du savoir ”²⁵, ils semblent omettre le lien unissant l’objectivité à la constitution intersubjective des significations des phénomènes naturels, sans laquelle on ne saurait pas rendre raison de l’objectivité en tant que *principe régulateur* de la recherche scientifique.

Et pourtant, la phénoménologie et l’épistémologie historique partagent le même projet dans la mesure où la constitution de l’objectivité représente, selon ces deux approches, un certain accomplissement, dont on peut retracer les étapes historiques en reconstituant les gestes constitutifs, “ nécessairement oubliés dès qu’accomplis ”,²⁶ ainsi qu’en reconduisant les opérations méthodiques d’idéalisation et de formalisation à leur origine, pour ne pas confondre la réalité avec ce qui n’était qu’une méthode de sa rationalisation.²⁷ L’histoire des techniques et des dispositions, acquises à travers les atlas, pourrait ainsi être complétée par l’analyse d’autres procédés d’objectivation, tels que la géométrisation des formes ou l’explication strictement causale, au moyen desquels la science s’achemine vers la précision, l’identité et la prédiction de tout phénomène naturel. La phénoménologie et l’épistémologie historique se révèlent ainsi complémentaires dans leur effort commun pour retracer la généalogie d’une signification objective des connaissances scientifiques. Qui plus est, aussi bien la phénoménologie que l’épistémologie historique s’opposent au naturalisme et à son objectivisme, qui contournent la tâche d’expliquer la genèse progressive des objets des sciences en lui préférant une construction toute métaphysique, consistant à postuler un rapport causal

doit avoir son index mathématique dans des événements du domaine des formes, conçu naturellement toujours comme déjà idéalisé ”.

²⁵ “ To be objective is to aspire to a knowledge that bears no trace of the knower – knowledge unmarked by prejudice or skill, fantasy or judgment, wishing or striving. Objectivity is blind sight, seeing without inference, interpretation, or intelligence. ” *Objectivity*, p. 17.

²⁶ Comme le dit Foucault à propos du partage entre la raison et la déraison. Voir Michel Foucault, *Dits et Ecrits*, Tome I, Paris, Gallimard, 2001, p. 189.

²⁷ Selon la thèse de Husserl, le modèle mathématique de la nature est confondu chez Galilée avec la nature elle-même. Voir *La crise des sciences européennes*, II, § 9 c – d.

entre une réalité physique, purement quantitative, et son image subjective dans notre conscience.

Katarina Lončarević
Faculty of Political Sciences
University of Belgrade, Serbia

Foucault and Epistemology Revisted: Could Genealogy Be Seen as Political Epistemology?*

I Foucault As an ‘Epistemologist’?

Michel Foucault (1926-1984) was interested in questions on knowledge and science from the very beginning of his intellectual career. However, his name and ideas are rarely mentioned in discussions within epistemology as a philosophical discipline. His case even represents a rare example of a consensus among analytical and continental philosophers, who mostly agree that Foucault is not an epistemologist.¹ Even Foucault himself would probably agree with both his critics and sympathizers when they say that he does not do epistemology, if ‘doing epistemology’ means providing a theory of knowledge in the usual sense of the word or offering an account of necessary and sufficient conditions for knowledge.²

In order to understand Foucault as an epistemologist, we need to redefine the term “epistemology”. In this paper, I will use Alcoff’s definition according to which epistemology should not be understood as a “foundationalist, anti-skeptical project. ... [E]pistemology [should be understood] as the theorizing about knowledge” in general.³

In Foucault’s writings, knowledge is not attributed to beliefs and/or sets of beliefs, and is not the sum of justified true beliefs. Knowledge includes much more: the space where the subject stands as a knower, and the field in which concepts are determined and

* A longer and somewhat different version of this paper was published in 2013. Cf. Katarina Lončarević, “Foucault’s Genealogy as Epistemology”, *Filozofski godišnjak*, vol. 26, 2013, pp. 75–96.

¹ Cf. Richard Rorty, “Foucault and Epistemology,” in: David Couzens Hoy ed., *Foucault: A Critical Reader*, Oxford, Basil Blackwell, 1986, pp. 41–50.

² Cf. James D. Faubion, “Introduction,” in James D. Faubion ed., *Essential Works of Foucault 1954-1984, Volume Two: Aesthetics, Method, and Epistemology*, New York, The New Press, 1998, p. xxiv.

³ Linda Alcoff, *Real Knowing: New Versions of the Coherence Theory*, Ithaca and London, Cornell University Press, 2008, p. 4.

statements are arranged.⁴ Foucault is against the conception of knowledge as a *possession* of *truth*, and against the conception of the self as the *unitary* condition of cognition. He is also against the conception of *truth* as an ultimate value of a reasoned inquiry.

At first, it looks like his works during the 1960s are the best candidates for an epistemological inquiry, and it is not surprising that when epistemologists read Foucault, which is rare, they usually read his ‘archaeological’ books. However, there are *at least* two important reasons why it is better to grasp Foucault’s critique of traditional epistemology in genealogical and not in archaeological works. Namely, after the events of May 1968, Foucault started to think more about the concept of *power*, which became the central concept of his genealogical writings, and second, at least in one essay he described genealogy as an ‘insurrection of subjugated knowledges’ and subjugated knowers, marginalized knowledges and knowers that could be seen as resistance against totalizing and dominant knowledges which eliminate others through violence.

During the genealogical period, Foucault never uses as examples of knowledge simple common-sense beliefs or perceptions – for example, “Grass is green” – or some beliefs from natural sciences. His examples come from ‘suspicious’ sciences, human and social ones, sciences with low epistemic status, and we might say that the starting point for Foucault’s ‘epistemological’ inquiries are not commonly accepted beliefs, but the most controversial and complicated beliefs from the sciences of man.

Foucault is primarily interested in understanding the production of knowledges and truths, in the plural, in their historical specificities. He wants to know how and why we think some things are true, and how and why we take some things as knowledge. His histories/genealogies are not meant to stay in history – they are not histories for their own sake. They are histories of the present, with an aim to show how our standards of knowledge have evolved, to explain their contingencies and to show that there is nothing necessary or inevitable in our dogmatic attachments with the current regimes of truth and knowledge.

⁴ Cf. Alcoff, *Real Knowing*, 144.

II The Genealogical Method

Foucault shares with Nietzsche the idea of genealogy as a *critique* – genealogy is a critique of the efforts to support established authorities on the basis of their *origin* (*Ursprung*). Traditional history is a project devoted to looking for ‘origins’ of things, discourses, and practices. According to Foucault, this search for origins tries to capture *the exact essence* of things, because it assumes the existence of immobile forms that precede the external world of accidents and succession. This search assumes and is directed to that which is *already* there. Contrary to traditional history, genealogy is an ‘effective history’, which means history that does not unify, but “distinguishes, separates, and disperses”⁵; a history that is opposed to the search for origins, or ahistorical foundations, essences and unchangeable truth. Genealogy is a history without constants, because contrary to the Cartesian tradition, it assumes that nothing in ‘man’ is stable enough to serve as a basis for self-recognition or for understanding others or the world. For a genealogist, there are no fixed essences, no underlying unchangeable laws. Genealogy reveals “the secret that [things] have no essences, or that their essence is fabricated”.⁶

In his genealogical works, Foucault under the influence of Nietzsche insists on the connection between power and knowledge. Foucault shares Nietzsche’s insight that knowledge is not outside the conflict; it is not at a safe distance from struggle which would enabled it to pronounce the truth of the conflict. There is no ‘outside’ of history, there is no neutral point from which it is possible to observe and proclaim the truth of various knowledge claims. Knowledge for both Nietzsche and Foucault is not a disinterested and objective observation, but a weapon: “knowledge is not made for understanding; it is made for cutting”.⁷

The changes in knowledge are not themselves the result and the product of thought itself. When knowledge contained in the sciences of man changes, the causes are the social forces that control the behavior of individuals. Although Foucault does not reduce knowledge to power, he is explicit that knowing is not the process that can achieve a total

⁵ Michel Foucault, “Nietzsche, Genealogy, History”, in *Essential Works of Foucault 1954-1984, Volume Two: Aesthetics, Method, and Epistemology*, ed. James D. Faubion, New York, The New Press, 1998, p. 380.

⁶ *Ibid.*, p. 371.

⁷ *Ibid.*, p. 380.

escape from power relations, and that power does not only have a constraining role, and does not operate only by eliminating some knowledge; it also has a positive epistemic role in producing knowledges, subjects and objects of knowledges and truths.

Throughout Foucault's genealogical period, he was influenced by Nietzsche's idea of looking for power behind the accepted and dominant knowledges, sciences, religions and other authorities, which usually present themselves as grounded in disinterested evidence and argument.

III Power

The key concept in Foucault's genealogical works is the concept of power. He does not provide a 'theory' of power. He does not offer a 'theory' of knowledge either. He usually speaks about what power is not, instead of providing a positive account of it. In that way, Foucault tries to show how his ideas on power *differ* from other proposed accounts, especially those given by liberal theorists and Marxists.

Both liberals and Marxists operate with a so-called "juridico-discursive"⁸ model of power that assumes that power is *possessed* (by the pre-social individual, the class or the people, etc.); that power is *centralized* (in law, in economy or in the state, etc.); and that power is primarily *repressive*. According to the 'juridico-discursive' model, power is essentially *negative* – it produces nothing but "limit and lack".⁹ Power is everywhere the *same* – it operates by endlessly reproduced mechanisms of law and censorship. Power is, finally, *domination* – all it can do is to forbid and command obedience.

Although liberals and Marxists do not have the same or even similar political and economic goals, they both consider that power acts in a negative and restrictive way towards its subjects. Precisely because power is negative and restrictive, it leaves space for potentially disobedient subjectivity, and that is the reason why the state and its power manifested through laws is important for both, liberals and Marxists. After the events of 1968, Foucault started to challenge explicitly this conception of power as negative and prohibitive, and to consider that account as too narrow and one-dimensional.

⁸ Michel Foucault, *The History of Sexuality: Volume One, An Introduction*, New York, Vintage Books, 1990, p. 82.

⁹ *Ibid.*, p. 83.

Foucault does not deny the ‘negative’ side of power, but believes that it is precisely that – one side, and therefore not a complete picture: we cannot understand power completely if we understand it as purely negative, as if it only says ‘no’. Negative conceptions of power have direct implications on the way some important epistemological concepts are understood. The Cartesian legacy and the liberal conception of state power as something negative result in very specific conceptions of knowledge and the knower. Traditional epistemology does not accept knowledge as a political event. If something is true knowledge, it has to be dissociated from power; it can appear only in isolated, free space of (meditating) subjectivity. This is the reason why Foucault in “Prison Talk” says that it is generally accepted that “[p]ower makes men mad, and [that] those who govern are blind; only those who keep their distance from power, who are in no way implicated in tyranny, shut up in their Cartesian *poêle*, their room, their meditations, only those can discover the truth”¹⁰.

Unlike liberals, Marxists do think that knowledge is political, and they even distinguish between two sorts of knowledge: the ‘scientific’ and ‘true’ knowledge that is in accordance with the dialectical understanding of history and ideology, or the distorted knowledge that is useless except for the ruling class, and therefore false. However, Marxists identify the truth with liberation and falsity (ideology) with subjection, which means that both liberals and Marxists believe in *one, objective, pure truth* and do not approve of its political critique.¹¹

The genealogical period in Foucault’s work brought a new conception of power and subsequently a new conception of knowledge. He explicitly rejects analyzing power within the paradigm of repression, because it cannot embrace all aspects of power: power is not only negative, but primarily creative. Power is not only prohibitive, rule-based and uniform, but first and foremost positive and productive; it is everywhere, omnipresent, dispersed all over the social body. As such, power *produces* and *creates* the knowing subjects. The relational nature of power depends on a multiplicity of points of resistance: if there is a relation of power, there is a possibility of resistance, which means for Foucault that we are never trapped by power – it is always possible to modify its hold. On the other

¹⁰ Michel Foucault, “Prison Talk,” in: Colin Gordon ed., *Power/Knowledge: Selected Interviews and Other Writings 1972-1977*, Brighton, The Harvester Press, 1980, p. 51.

¹¹ Todd May, *Between Genealogy and Epistemology: Psychology, Politics and Knowledge in the Thought of Michel Foucault*, Pennsylvania, The Pennsylvania University Press, 1993, p. 38.

hand, Foucault in his later writings also emphasizes that power relations can be implemented only where there is a possibility of resistance, that is in cases where there is a struggle, a conflict, a dynamical relationship between different actions with the possibility of different outcomes. There are no power relationships where there is no possibility of resistance and modification of current temporary alignments of local forms of power. To sum up, genealogy sees power as a productive and creative network of relations of forces that produces and shapes knowledges and objects of knowledge, and utilizes the effects of knowledges.

IV Concepts of Knowledge and the Subject of Knowledge

By the beginning of the 1970s, Foucault adopted Nietzsche's view that knowledge itself is a product of struggles for power and domination. Throughout the genealogical period, Foucault rarely spoke about the concept of knowledge in general and was not inclined to give a general account of anything including epistemological problems. In this section, I argue a little bit contra Foucault. Namely, I see the first 'methodological' lecture of the series he gave in May 1973 in Brazil, published under the title "Truth and Juridical Forms", as his attempt to explain the concepts of knowledge and of the subject of knowledge on a more general level, but differently that it is done in traditional epistemology:

"Two or three centuries ago, Western philosophy postulated, explicitly or implicitly, the subject as the foundation, as the central core of all knowledge, as that in which and on the basis of which freedom revealed itself and truth could blossom. ... [I]n the field of what we may call the 'theory of knowledge', or in that of epistemology, ... it seems to me that the theory of the subject has remained very philosophical, very Cartesian and Kantian."¹²

So, today when someone tries to do the history of knowledge, he or she "sticks to this subject of knowledge, to the subject of representation as the point of origin from which knowledge is possible and truth appears."¹³

¹² Michel Foucault, "Truth and Juridical Forms," in: James D. Faubion ed., *Essential Works of Foucault 1954-1984, Volume Three: Power*, New York, The New Press, 2000, p. 3.

¹³ *Ibid.*

Foucault's project is to do something completely different. Using his genealogical method, he wants to see how a subject comes to be *constituted*, that is not definitively given, how the subject constitutes itself *within*, not outside history, and how the subject is constantly established and reestablished by history as different and variable. In order to do that, he needs to do something else first – to investigate the concept of knowledge.

Foucault finds the source for the new reconception of knowledge in one of Nietzsche's early essays, "On Truth and Lie in an Extra-moral Sense". At the very beginning of the essay Nietzsche says that knowledge is an invention of clever animals.¹⁴ Invention is on the one side a break, but on the other it is something with a small beginning – low, mean, and Foucault wants to counterpose these small and low beginnings of fabrications to the solemnity of origins.¹⁵

"To say that [knowledge] was invented is to say that it has no origin. More precisely, it is to say, however paradoxical this may be, that knowledge is absolutely not inscribed in human nature. Knowledge doesn't constitute man's oldest instinct; and, conversely, in human behavior, ... the human instinct, there is no such thing as a seed of knowledge."¹⁶

At this point, Foucault detaches himself from the philosophical tradition in a very radical way. First, he does not start his reflections about knowledge with, for example, the opening lines of Aristotle's *Metaphysics*: "ALL men by nature desire to know".¹⁷ Instead, Foucault quotes the opening lines of an essay written by a 'suspicious' figure in the history of philosophy, Nietzsche, and claims, contrary to the whole tradition starting from Plato and Aristotle, that knowledge "is absolutely not inscribed in human nature".

Knowledge has connections with instincts, but it cannot be present in them, and it is not one instinct among the others. Knowledge has a basis in instincts, but the basis is the *confrontation*, the struggle between instincts. Knowledge arises as an outcome of a combat, and it is "not instinctive, it is counterinstinctive; just as it is not natural, but counternatural".¹⁸ And this is the first sense in which Foucault speaks about knowledge as an 'invention', and not as something that is derived from human nature.

¹⁴ Cf. Friedrich Nietzsche, *The Portable Nietzsche*, New York, Penguin, 1976.

¹⁵ Where traditional historians see continuities, origins and depth, the genealogist sees discontinuities, surfaces of events, small details and minor shifts. That is the reason why Foucault speaks about "the small beginning" of the invention of knowledge.

¹⁶ *Ibid.*, pp. 7–8.

¹⁷ Aristotle, *Metaphysics*, London, NuVision Publication, 2005, p. 7.

¹⁸ Foucault, "Truth and Juridical Forms", p. 8.

There is a second meaning of ‘knowledge as an invention’. Namely, knowledge “isn’t even closely connected to the world to be known”.¹⁹ In other words, the *novelty* of knowledge means that knowledge is not just an expression of our experience of things in the world. There is no resemblance or prior affinity between knowledge and things to be known. Therefore, Foucault claims: “In ... strictly Kantian terms, one should say the conditions of experience and the conditions of the object of experience are completely heterogeneous.”²⁰

Knowledge is an attempt to *impose* the order in a chaotic world. It is not just the attempt to describe things in the world, because knowledge is not a description.²¹ “Knowledge can only be a violation of the things to be known, and not a perception, a recognition, an identification of or with those things.”²² What knowledge does is imposing on things in the world “an order which is new”,²³ and the relation between knowledge and things to be known is *never* merely representational, because there is no relation of continuity between knowledge and things to be known, but the relations of power and struggle, which means that we do not have to postulate (like Descartes had to) the existence of God, who traditionally has guaranteed the continuity and harmony between knowledge and the things in the world. However, if there is not relation of continuity between knowledge and instincts, but only discontinuity and power relations, “then it’s not God that disappears, but subject in its unity and sovereignty”.²⁴ If the mechanisms of instincts are on a completely different level than knowledge is, we do not need to postulate the unity of human subject that has been traditionally given beforehand and as definite.

Foucault follows Nietzsche who argues that there is no nature of knowledge, no essence of knowledge and no universal conditions of knowledge.

“[K]nowledge is always the historical and circumstantial result of conditions outside the domain of knowledge. In reality, knowledge is an event that falls under the category of activity. Knowledge is not a faculty or a universal structure. ... [K]nowledge ... only belongs to the order of results, events, effects.”²⁵

¹⁹ *Ibid.*

²⁰ *Ibid.*, p. 9.

²¹ Cf. Michael Kelly, *The Political Philosophy of Michel Foucault*, New York – London, Routledge, 2008, p. 19.

²² Foucault, “Truth and Juridical Forms,” p. 9.

²³ Kelly, *The Political Philosophy of Michel Foucault*, p. 20.

²⁴ Foucault, “Truth and Juridical Forms”, p. 10.

²⁵ *Ibid.*, pp. 13–14.

This Nietzschean view of knowledge as a product, result and effect of struggles for power that Foucault adopts, has important consequences for the understanding of the 'subject' of knowledge. Foucault does not understand the subject as universal, timeless or abstract, the source of how one makes sense of the world, the foundation of all knowledge, thought and action. For Foucault, the subject of knowledge/knower traditional epistemology speaks about does not exist in his/her autonomy and universality. Foucault explicitly rejects the subject of the Enlightenment understood as an *a priori* subject of knowledge: "What I refused was precisely that you first of all set up a theory of the subject. ... What I wanted to know was how the subject constituted himself in such and such determinate form".²⁶ The subject is an effect, the product of specific power and knowledge constellation. That subject is not prior to history, and not pre-given. It is created and changed by outside events; it is constantly dissolved and recreated in different configurations along with other forms of knowledge and social practices.²⁷ This conception of the subject of knowledge as an effect of power and knowledge constellations, or this dismissal of the traditional knower as one of the central epistemological categories, is probably the most radical Foucault's epistemological move in his genealogical works. To put it simply, epistemology is not based on the concept of the knower, and knowledge does not have a cause in an independently existing knower, opposed to the world and to other knowers.

Foucault thus rejects the constituting knowing subject of the Enlightenment's epistemology. However, this rejection, worries some philosophers, who argue that we need the notion of that subject in order to have political and moral agency.²⁸ Alcoff, for example, worries that, according to Foucault, we, as subjects, are *overdetermined*, constructed by power/knowledge relations, social discourse and cultural practice.²⁹ Foucault's subject is, therefore, "totally determined, because it is enmeshed in relationships of power and is produced as effect through disciplines and practices",³⁰ and unable to resist. However, Foucault does not completely reject or abandon the subject.

²⁶ Foucault quoted in Margaret McLaren, "Foucault and the Subject of Feminism," *Social Theory and Practice*, vol. 23, n° 1, 1997, p. 112.

²⁷ Cf. Clare O'Farrell, *Michel Foucault* (London: Sage Publications, 2005), pp. 110–113.

²⁸ Cf. McLaren, "Foucault and the Subject of Feminism," p. 110.

²⁹ Linda Alcoff, "Cultural Feminism Versus Post-Structuralism: The Identity Crisis in Feminist Theory," *Signs*, vol. 13, n° 13, 1988, p. 416.

³⁰ McLaren, "Foucault and the Subject of Feminism," p. 110.

Instead, he rejects the ‘philosophy of the subject’, the idea of One, universal, disembodied subject, out of space and time and outside power relations.

Foucault’s conception of the subject of knowledge displaces the traditional dichotomy between the constituting Cartesian subject, who possesses agency and autonomy, and the constituted subject that is wholly determined by social forces. For Foucault, the subject is constituted but it is at the same time an agonism, a permanent provocation to power/knowledge constellation that defines its subjectivity.³¹ Only within the Enlightenment framework that is based on dualisms, such as subject/object, rational/irrational, etc., the constituted subject is passive and without agency. In Foucault’s work, where the dualist thinking is rejected, the constituted subject is the subject who acts and resists.

In other words, it can be argued that the Foucauldian approach to the subject of knowledge rejects the epistemology that rests on traditional dualisms, and therefore, could escape objections regarding the possibility of action in the domain of the political and the possibility of knowledge in the domain of epistemology.

V Power and Knowledge: Power-Knowledge

The picture of a genealogy of knowledge that Foucault provides is more than problematic both for Foucault’s critics and opponents, and for his sympathizers. Allegedly, Foucault’s genealogy reduces knowledge to the struggle and the play of forces, or to be more specific, he identifies knowledge with power.

Foucault’s scholars, in an effort to defend him from critiques that the only operative force in his account of genealogy of knowledge is power, often quote the famous passage from *Discipline and Punish*, as the clearest Foucault’s general account of the relationship between power and knowledge, although he tried on various occasions, especially in interviews he gave, to defend himself from the accusation for the reduction of knowledge to power. In one of those interviews, Foucault explicitly says:

“[W]hen I read – and I know it was being attributed to me – the thesis ‘Knowledge is power’ or ‘Power is knowledge’, I begin to laugh, since

³¹ Susan Hekman, *Gender and Knowledge: Elements of a Postmodern Feminism*, Cambridge, Polity Press, 1990, pp. 72–73.

studying their *relation* is precisely my problem. If they were identical, I would not have to study them and I would be spared a lot of fatigue as a result. The very fact that I pose the question of their relation proves clearly that I do not *identify* them.”³²

However, this is not understood as ‘enough’, or as a sufficient explanation by many of Foucault’s critics,³³ and therefore, it is urgent to explore Foucault’s general account of the *relation* between knowledge and power he gave in *Discipline and Punish*:

“Perhaps, ... we should abandon a whole tradition that allows us to imagine that knowledge can exist only where the power relations are suspended and that knowledge can develop only outside its injunctions, its demands, and its interests. Perhaps we should abandon the belief that power makes mad and that, by the same token, the renunciation of power is one of the conditions of knowledge. We should rather admit that power produces knowledge (not simply by encouraging it because it serves power or by applying it because it is useful); that power and knowledge directly imply one another; that there is no power relation without the correlative constitution of a field of knowledge, nor any knowledge that does not presuppose and constitute at the same time power relations. These ‘power-knowledge relations’ are to be analysed, therefore, not on the basis of a subject of knowledge who is or is not free in relation to the power system, but, on the contrary, the subject who knows, the objects to be known, and the modalities of knowledge must be regarded as so many effects of these fundamental implications of power-knowledge and their historical transformations. In short, it is not the activity of the subject of knowledge that produces a corpus of knowledge, useful or resistant to power, but it is power-knowledge, the processes and struggles that traverse it and of which it is made up, that determines the forms and possible domains of knowledge.”³⁴

Foucault here speaks of knowledge and power as, on the one side, heterogeneous, but on the other, as mutually reinforcing. Power provides the site for the elaboration of knowledge, and knowledge itself has a constituting effect on power relations in, for example, establishing or reinforcing the hierarchies of epistemic authority. The passage also speaks about differences between the traditional view of knowledge and power and the genealogical view. In traditional accounts, knowledge is possible only in the absence

³² Michel Foucault, “Structuralism and Post-structuralism”, in *Essential Works of Foucault 1954-1984*, Volume Two, p. 455.

³³ Those critics are different philosophers and political scientists from Richard Rorty to Jurgen Habermas and Charles Taylor.

³⁴ Michel Foucault, *Discipline and Punish: The Birth of the Prison*, New York, Vintage Books, pp. 27–28.

of power relations, and the subject of knowledge comes first, or as Foucault says, traditional accounts, including liberal and Marxist ones, are based on “the underlying idea that relations of force, economic conditions, and social relations ... are imposed on the subject of knowledge that remains identical”.³⁵ The genealogical view, on the other hand, offers the picture of knowledge that is a product of power-knowledge relations, and the picture of the subject of knowledge that comes later as the product, the effect, and not the source or the cause of power-knowledge relationships.³⁶

Knowledge and power are both part of the system that creates the conditions for the production and authorization of new knowledges, modes of knowing and objects of inquiry. Neither knowledge nor power can be reduced to or replaced with one another in discursive analysis. ... [K]nowledge has no autonomous existence apart from power.³⁷

Therefore, Foucault speaks about reciprocal, interconnected and interdependent relation of power and knowledge, and he does not claim that epistemology should disappear or be reduced to political analysis. He claims that there cannot be an epistemological analysis that is autonomous and separated from political analysis.

VI Knowledge As Resistance: Genealogy As an Insurrection of Subjugated Knowledges

Although Foucault ‘defines’ power as depending on resistance, and although he claims that power always implies resistance, resistance is the concept that is underdeveloped in Foucault’s work.³⁸ In the lecture Foucault gave in 1976, he explicitly speaks about ‘subjugated knowledges’ and their potential to be forms of resistance to dominant knowledge/knowledges, and about genealogy as a method of resistance.

Although Foucault in many of his interviews tries to clarify his view on the concept of resistance, I want to argue that the first lecture in the series of *Society Must Be Defended* is very important, because there we can see in the most precise way how knowledge/knowledges can be seen as tools of resistance.

³⁵ Foucault, “Truth and Juridical Forms,” p. 15.

³⁶ Cf. May, *Between Genealogy and Epistemology*, p. 72; Alcoff, *Real Knowing*, p. 153.

³⁷ Alcoff, *ibid.*

³⁸ McLaren, “Foucault and the Subject of Feminism,” p. 123.

A genealogy of knowledges explains the *struggle* of one knowledge or *episteme* against other knowledges or *epistemes*. The epistemologically important thing here is that what we currently call ‘knowledge’ is actually one knowledge, the dominant one, the knowledge that in a struggle defeated other knowledges. On the other hand, genealogy as a method is designed to facilitate “the insurrection of subjugated knowledges”.³⁹ Under the term ‘subjugated knowledges’, Foucault means two things. First, the term refers to “blocks of historical knowledges that were present in the functional and scientific ensembles, but which were masked”.⁴⁰ The genealogist, then, reactivates them and reveals their existence. Second, and more importantly, the term refers to

“a whole series of knowledges, as insufficiently elaborated knowledges: naïve knowledges, inferior knowledges, knowledges that are below the required level of erudition or scientificity. ... [T]he knowledge of ... the patient, the nurse, the doctor, that is parallel to, marginal to, medical knowledge, the knowledge of the delinquent, ... [knowledge of] what people know (and this is by no means the same thing as common knowledge or common sense but, on the contrary, a particular knowledge, a knowledge that is local, regional, or differential, incapable of unanimity and which derives its power solely from the fact that it is different from all knowledges that surround it), ... at a local level, ... disqualified knowledges.”⁴¹

Therefore, genealogy uses history to give a *voice* to the marginal and submerged knowledges that are ‘a little beneath history’. Foucault wants to show that some paths in the history were not taken and that there is a number of unactualized possibilities and events that do not fit the dominant functionalist schema of knowledge in sciences of man.

Genealogy is about the insurrection or rebellion of knowledges not against the scientific methods, concepts or content, but against “the centralizing power-effects that are bound up with the institutionalization and workings of any scientific discourse organized in a society such as ours”.⁴² This means that when something is claimed to be a ‘science’ or something has an aspiration to be a ‘science’, Foucault in that claim or aspiration does not see the rational structure of that potential ‘science’, or that its propositions are rationally verified. On the contrary, he sees in that kind of claim or

³⁹ Michel Foucault, *Society Must Be Defended: Lectures at the College de France 1975-1976*, London, Penguin, 2003, p. 7.

⁴⁰ *Ibid.*

⁴¹ *Ibid.*, pp. 7–8.

⁴² *Ibid.*, p. 9. By ‘society such as ours’ Foucault always thinks of the Western societies.

aspiration an attempt to *disqualify* some other knowledges, and to *attach* itself to scientific discourse.⁴³ Genealogy, as a method, is therefore, an attempt

“to desubjugate historical knowledges, to set them free, or in other words, to enable them to oppose and struggle against the coercion of a unitary, formal, and scientific theoretical discourse. The project of these disorderly and tattered genealogies is to reactivate local knowledges ... against the scientific hierarchicalization of knowledge and its intrinsic power-effects.”⁴⁴

Why is Foucault’s idea of genealogy as the insurrection of subjugated, oppressed and local knowledges epistemologically relevant? First of all, Foucault’s account of knowledges in the plural, dissociates the concept of knowledge from the concept of truth. Traditionally speaking, knowledge cannot be knowledge if it is not true; there cannot be a false knowledge. And the traditional belief that “We all agree that what we know is true” gains a completely new meaning in Foucault’s account of the plurality of knowledges, of their mutual struggle and of one of them ‘winning’ and becoming the knowledge that dominates over a number of different, local knowledges. What ‘we all’ allegedly agree in our societies is that the one, dominant knowledge, or what we call ‘science’, is ‘true’, and that other oppressed, subjugated, local and naïve knowledges are ‘false’. They are false, not because they are ‘irrational’ or not in accordance with some ‘objective’ methods of inquiry (if there are such methods). They are false because they cannot fit into the mold of the dominant scientific and epistemological discourse. Consequently, some of them in some other constellation of power relations could be true.

Second, subjugated or disqualified knowledges can subvert or undermine the supremacy of dominant and ‘hegemonic’ knowledge(s). But, this is still not enough for a radical critique. Namely, what Foucault sees in dominant, totalizing knowledges is that they can work or operate only if they exercise some kind of *coercion* and *violence* against local and particular knowledges, because they *need* to subsume them under their universal and totalizing structure.⁴⁵ Linda Alcoff sees Foucault’s account of subjugated knowledges as important because they are, according to her,

“valorized not because they represent more accurate representation of the Real as it exists in itself, ... but because they do not require the amount of violence,

⁴³ Foucault, *Society Must Be Defended*, p. 10.

⁴⁴ *Ibid.*

⁴⁵ Cf. Alcoff, *Real Knowing*, p. 154.

distortion, and omission that global knowledges require. ... [In other words], they have a *different* relationship to power, and ... this different relationship will constitute a different field [of knowledge].”⁴⁶

The structure or the form of subjugated and local knowledges is *different*, because they have a different position in relation to the dominant systems of power. They are not local or subjugated because they have not reached domination yet, but because “they do not aspire to dominance. They do not construct competing unitary, formal, totalizing theoretical systems.”⁴⁷ Linda Alcoff, in her account of Foucault’s idea of ‘subjugated knowledges’ argues that the important distinction that should be made here is not the one between dominant and subjugated knowledges, which Foucault actually makes, but the one between *local* and ‘*hegemony-seeking*’⁴⁸ knowledges, because the second distinction allows us to mark important differences that exist between the so-called ‘subjugated knowledges’, and to open the possibility to think about epistemology in new, productive, innovative ways.⁴⁹

First, although it can be argued that there is a large number of subjugated knowledges, including knowledges of, for example, Neo-Nazis groups, there is a criterion Foucault could use and according to which knowledges of those groups cannot be considered as places for critical interventions. The main question/criterion is whether those kinds of ‘knowledges’ seek for disqualification and marginalization of other knowledges. In the case of Neo-Nazis groups, it is more than obvious that their knowledge seeks domination and requires, or even includes the exercise of violence against other knowledges, as well as their disqualification.

The second question that should be addressed is: How Foucault can be sure that his critique does not lead towards the mere turning of the subjugated knowledge into the dominant position? Although Foucault would probably answer that we can never be sure and that we cannot predict the future, again, the distinction between local and ‘hegemony’-seeking knowledges could help: the mere reversal of knowledges will happen only in the case when the subjugated knowledge is a ‘hegemony’-seeking one. In the case when subjugated knowledge is not ‘hegemony’-seeking, there is a possibility for a new relation

⁴⁶ *Ibid.*, p. 155.

⁴⁷ *Ibid.*, p. 156.

⁴⁸ In this distinction, Alcoff uses the Marxist term ‘hegemony’, which Foucault does not use.

⁴⁹ Cf. Alcoff, *Real Knowing*, p. 157.

between knowledge and power, and for a new, different kind of knowledge than the one represented in ‘science’ or in dominant knowledge in our societies.

The important consequence of Foucault’s account of subjugated knowledges in the first lecture of *Society Must Be Defended*, is that subjugated, local knowledges can be a form of a *resistance* to dominant knowledges and discourses. Precisely because subjugated knowledges are local and spatially/temporally situated, and precisely because they are *different* from the dominant ones, they are also ‘innovative’ and ‘creative’, and they can be a form of resistance.⁵⁰

Through the account of subjugated or local knowledges, Foucault actually gives us the possibility to imagine a *different* kind of epistemology not as a solution for our current epistemological issues and theories that can replace or overturn the dominant paradigm as a better or more correct way of describing the world. Namely, what Foucault’s general epistemological reflections show is that we should focus on *specific, local and particular*, because we cannot escape, despite all the efforts, our temporal, cultural, political, and local specificity and particularity. This leads some theorists, such as Linda Alcoff, to claim that “[t]he specificity and locatedness of epistemology ... is perhaps the most important distinction to be made between [Foucault’s] conception of epistemology and traditional accounts.”⁵¹

I would like to argue, therefore, that the genealogical method can be seen as epistemological, and that it can help us in detaching ourselves from the dogmatic, transcendent epistemological formulations and can create the possibility for us to imagine a new, different epistemological project. “Providing [the genealogies] of present day ‘absolute’ truths will have a liberating effect, dislodging their power and thus freeing us to imagine new possibilities.”⁵²

To sum up, Foucault in his genealogical period, under Nietzsche’s influence, develops a critique of traditional epistemology and its core concepts of knowledge and of the knowing subject, that were imagined as dissociated from everything political, local, social and culturally specific. He argues in favor of marginalized voices and knowledges, because he sees them as potential sources of resistance against dominant, totalizing

⁵⁰ For Foucault’s account of the relationship between power and resistance, cf. Foucault, *The History of Sexuality*, p. 95–96.

⁵¹ Alcoff, *Real Knowing*, p. 160.

⁵² *Ibid.*, p. 119.

knowledges that deny their situatedness and historical specificity, proclaiming themselves to be absolute and universal truths.

VII Conclusion

The interpretation of Foucault's genealogical method as epistemological has shown that genealogy is a radical *critique* of knowledge, science and truth as they appear within the traditional epistemological framework. I have argued that Foucault's genealogy should be seen as one of the biggest challenges to our uncritical commitments to allegedly neutral, objective and ahistorical reasoning. Foucault, following Nietzsche, breaks with the traditional understanding of the subject as a foundation of knowledge, of knowledge as disinterested and apolitical event, and as something that should be kept at a safe distance from power. Because power in Foucault's account is primarily productive and creative, Foucault provides an understanding of knowledge and of the so-called subject of knowledge, that is completely different than the traditional one. For Foucault, knowledge is always situated, local and contextual. Knowledge is an event, an effect, not a faculty or a universal structure. The subject/knower is not *unchangeable*; it is always produced and reproduced, established and reestablished in its very form through history. Foucault provides an 'epistemology' without the knowing subject in the traditional sense, and this move is the most radical moment in his epistemological work.

Generally speaking, Foucault's contribution to epistemology consists precisely in his insistence that our theories should be focused on the specific, the local and the particular; that our theories are not meant to stay with us forever, that they are changeable, dynamic and temporary, and that we should use them as tools in specific situations that we see as occasions for intervention and attack. On the other hand, genealogy as a critical method could help us to see potential dangers in our theoretical commitments and attachments. Finally, what Foucault has realized is that we cannot get out of power relations, and that there is no 'innocent' position from which we can be sure that what we are doing is right and that our theories are purely 'liberatory'. The genealogical critique reminds us of the possible dangers of our accounts and prevents us from becoming too comfortable with our own position and 'truths'.

References

- Alcoff, Linda. "Cultural Feminism Versus Post-Structuralism: The Identity Crisis in Feminist Theory." *Signs* 13/3 (1988): 405–436.
- Alcoff, Linda. *Real Knowing: New Versions of the Coherence Theory*. Ithaca and London: Cornell University Press, 2008.
- Aristotle. *Metaphysics*. NuVision Publication, 2005.
- Dreyfus, Hubert L., and Paul Rabinow. *Michel Foucault: Beyond Structuralism and Hermeneutics*. Harvester Wheatsheaf, 1982.
- Faubion, James. D. "Introduction." In *Essential Works of Foucault 1954-1984, Volume Two: Aesthetics, Method, and Epistemology*, edited by James D. Faubion, xiii–xliii. New York: The New Press, 1998.
- Foucault, Michel. "Prison Talk." In *Power/Knowledge: Selected Interviews and Other Writings 1972-1977*, edited by Colin Gordon, 37–54. The Harvester Press, 1980.
- Foucault, Michel. *The History of Sexuality: Volume One, An Introduction*. New York: Vintage Books, 1990.
- Foucault, Michel. *Discipline and Punish: The Birth of the Prison*. New York: Vintage Books, 1995.
- Foucault, Michel. "Nietzsche, Genealogy, History." In *Essential Works of Foucault 1954-1984, Volume Two: Aesthetics, Method, and Epistemology*, edited by James D. Faubion, 369–391. New York: The New Press, 1998a.
- Foucault, Michel. "Structuralism and Post-structuralism." In *Essential Works of Foucault 1954-1984, Volume Two: Aesthetics, Method, and Epistemology*, edited by James D. Faubion, 433–458. New York: The New Press, 1998b.
- Foucault, Michel. "Truth and Juridical Forms." In *Essential Works of Foucault 1954-1984, Volume Three: Power*, edited by James D. Faubion, 1–89. New York: The New Press, 2000.
- Foucault, Michel. *Society Must Be Defended: Lectures at the College de France 1975-1976*. London: Penguin, 2003.
- Hekman, Susan. *Gender and Knowledge: Elements of a Postmodern Feminism*. Cambridge: Polity Press, 1990.
- Kelly, Michael. *The Political Philosophy of Michel Foucault*. New York, London: Routledge, 2008.
- Lončarević, Katarina. "Foucault's Genealogy as Epistemology." *Filozofski godišnjak* 26 (2013): 75–96.
- May, Todd. *Between Genealogy and Epistemology: Psychology, Politics and Knowledge in the Thought of Michel Foucault*. University Park, Pennsylvania: The Pennsylvania University Press, 1993.
- McLaren, Margaret. "Foucault and the Subject of Feminism." *Social Theory and Practice* 23/1 (1997): 109–127.

- Nietzsche, Friedrich. *The Portable Nietzsche*. New York: Penguin, 1976.
- O'Farrell, Clare. *Michel Foucault*. London: Sage Publications, 2005.
- Rorty, Richard. "Foucault and Epistemology." In *Foucault: A Critical Reader*, edited by David Couzens Hoy, 41–50. Oxford: Basil Blackwell, 1986.

Une épistémologie du « cas clinique » : l'exemple de l'aliénisme de Pinel¹

Introduction

Dans cet article, j'entends analyser la spécificité du cas clinique tel qu'il apparaît dans l'aliénisme de Pinel, et comment la structure de son récit éclaire certains aspects de l'institution de la psychiatrie médicale.

Le cas clinique est si naturellement vu comme un objet de plein droit médical, qu'il nous semble que le médecin parle de cas comme le botaniste parle de plantes. Rien de plus simple : un cas est une petite histoire racontant comment quelqu'un tombe malade, de quelle maladie, ce qui lui arrive ensuite, comment on le traite, et enfin s'il s'en sort. Alors, "cas" désigne aussi bien une catégorie ontologique - un cas est une entité spatio-temporelle singulière, par opposition à une classe, une propriété ou une règle - qu'une catégorie discursive, le cas s'opposant à la démonstration, à l'inférence ou à la taxinomie. Entité médicale par excellence, le cas semble certes ne pas avoir attendu la médecine dite clinique dont l'heure de gloire fut l'école de Paris du tournant du 19^{ème} siècle (Laennec, Corvisart, Bichat, etc.). Aristote le disait déjà, la médecine ne soigne pas l'homme, mais cet homme-ci : elle a vocation à l'individuel. Il n'y a pas de "cas" botanique ou physique, le cas est essentiellement médical car la médecine est essentiellement individuelle (d'où son statut toujours débattu, entre science – discours du général - et art – pratique du singulier – sur le fond de l'interdit aristotélicien d'une science de l'individuel). Le cas, histoire singulière par essence, n'a ainsi de sens qu'à s'écarter de la norme, laquelle pourrait s'énoncer a

¹ Ce texte est initialement paru dans la revue *Philosophie* (2013, vol. 3, p. 67-94) sous le titre: "Ecrire le cas-Pinel aliéniste". Merci à la revue et à ses rédacteurs en chef pour l'autorisation de le reproduire ici, minimalement modifié.

priori, comme “ santé ”.² En même temps, le cas a rapport à un universel, au sens où la maladie dont il est un cas est, elle, un concept universel. Enfin, le cas dans un écrit médical est singulier en un autre sens : il expose une pratique thérapeutique singulière : “ moi, Celse, moi Sydenham, ou moi, Cullen, j’ai diagnostiqué en cet individu telle maladie universelle, et l’ai traitée de telle manière et il a guéri (ou pas....) ”.

Le cas en médecine institue alors une double articulation : entre l’universel – telle ou telle pathologie – et l’individuel – tel individu malade et son histoire; entre l’ordre du visible constitué par les symptômes, et l’ordre invisible des signifiés de ces symptômes, comme de l’étiologie qui doit expliquer cet invisible. La configuration de la médecine, en une époque donnée, se reconnaît et se comprend aussi à la manière dont l’usage du cas clinique effectue cette double articulation. Ainsi, le discours sur le cas est rien moins que naturel, il obéit au contraire à certaines exigences et épouse certaines formes qui sont propres à différents périodes de l’histoire médicale, différentes disciplines, différentes écoles, qui définissent ce qu’on pourrait appeler le régime du cas. Le présent article vise à montrer comment avec le *Traité médico-philosophique* l’aliénisme donne à la double articulation qu’est un cas en médecine un tour spécifique dont la marque imprimera ultérieurement la psychiatrie, autrement dit à esquisser le régime du cas aliéniste, ou du moins pinellien.

Le *Traité médico-philosophique sur l’aliénation mentale ou la manie* est le livre majeur de l’aliénisme, celui dans lequel le mythe de la naissance de la psychiatrie rationnelle voit la rupture majeure avec les pratiques médicales anciennes et barbares de traitement des fous, et celui qu’inversement les démystificateurs à la suite de *l’Histoire de la Folie*, Foucault n’ont eu de cesse de replacer dans la trame d’un mouvement occidental de domestication de la déraison. Pinel y définit le champ de l’aliéniste, c’est-à-dire d’une médecine des troubles mentaux ayant sa légitimité propre, son objet - la maladie mentale, identifiée comme manie, en opposition avec les maladies somatiques mais aussi avec les fièvres délirantes - et enfin une méthode caractéristique, le “ traitement moral”. La signification épistémologique et le rôle historique du *Traité médico-philosophique* ont été

² Le langage quotidien l’atteste, qui dit “ c’est un cas ” lorsque quelqu’un semble bien loin de la normale.

longuement analysés depuis plusieurs décennies.³ Tout lecteur de cet ouvrage (en particulier dans sa première édition, étudiée ici, de 1801 (an IX)) remarquera qu'en grande partie il est constitué d'une collection de cas plus ou moins longs. Contrastant avec l'ambition théorique ultérieure d'Esquirol, dont le *Des maladies mentales* (1818) place simplement des éléments cliniques comme illustration de sa définition des entités théoriques et de sa discussion des classifications, le *Traité* pinellien est entièrement construit sur des cas cliniques, qui parfois tiennent lieu du développement théorique. Les cas distinguent d'ailleurs ce traité des communications faites auparavant par Pinel à l'Institut, qui établissaient des tableaux statistiques sur les entrées, les guérisons, etc.⁴; ces Mémoires font certes le corps du traité, mais Pinel les a réinvestis dans un ouvrage où le "cas clinique" tient une part majeure. Dans cette mesure, comprendre ce qui s'est joué avec l'inauguration de l'aliénisme implique de se demander à un certain point pourquoi cela a eu lieu sous la forme, non d'une nosologie ou d'un traité théorique, mais bien d'une collection de cas commentés.

Le cas clinique dans le *Traité* marque une double différence : d'une part entre le régime du cas à l'époque de la médecine clinique, et le régime du cas médical antérieur ; de l'autre entre l'usage du cas dans le *Traité* pinellien, et le cas dans la médecine clinique usuelle. Pour cette seconde différence, le travail de Bichat, contemporain du *Traité*, nous servira de premier repère ; un second repère sera fourni par la *Nosographie philosophique* de Pinel lui-même, dont il sera intéressant de constater que l'usage et la grammaire du cas diffèrent entement de ce que l'on y trouve dans le *Traité*.

Dans un premier temps, j'analyserai donc le contexte de cette double différence, qui situe le régime du cas pinellien. Ce sera d'abord le contexte historique institutionnel : l'hôpital ; je retracerai brièvement ses effets sur le régime du cas médical. Ensuite, il s'agit d'un contexte proprement philosophique, à savoir la référence commune de Pinel, Cabanis

³ Voir Dora Weiner, "Observe and Heal : Philippe Pinel's Experiment at the Salpêtrière Hospice, 1802–1805", in E. J. Engstrom, M. M. Weber, and P. Hoff eds., *Knowledge and Power : Perspectives in the History of Psychiatry*, Berlin, Verlag für Wissenschaft und Bildung, 1999 ; Dora Weiner, *Comprendre et soigner*, Paris, Fayard, 2001 ; Gladys Swain et Marcel Gauchet, *La pratique de l'esprit humain*, Paris, Gallimard, 1995 ; Gladys Swain, *Le sujet de la folie*, Paris, Payot, 1978 ; Jackie Pigeaud, *Aux portes de la psychiatrie. Pinel, l'ancien et le moderne*, Paris, Flammarion, 2002 ; Michel Foucault, *Histoire de la folie à l'âge classique*, Paris, Gallimard, 1961.

⁴ Textes publiés et commentés dans Dora Weiner "Observe and Heal : Philippe Pinel's Experiment at the Salpêtrière Hospice, 1802–1805", et Dora Weiner, *Comprendre et soigner*.

et Bichat : les Idéologues, eux-mêmes souvent médecins, et le sillage de la pensée de Cabanis.

Dans un second temps j'analyserai le régime du cas psychiatrique dans le texte de Pinel lui-même. L'analyse ne pourra pas éviter de sélectionner un nombre significatif de cas complets, afin d'en interroger la grammaire. J'en ai découpé sept dans le texte. Après avoir analysé quelques points généraux, je considérerai successivement trois aspects : le type de temporalité que déploient ces cas, ainsi que leur rapport à la causalité ; la présence de l'hôpital comme lieu de guérison ; la fonction rhétorique du cas dans l'ensemble du projet pinellien. Ressaisis selon ces trois dimensions, j'en conclurai que la multitude des recensions de cas et leur facture spécifique, dans le *Traité*, n'est point anecdotique mais participe au contraire du mouvement même de ce qu'on pourrait appeler la démonstration pinellienne – soit, démontrer que la folie est une maladie, donc comme telle curable, qu'il en existe une médecine spécifique, relativement autonome par rapport à la médecine établie, et que cette médecine a son lieu naturel dans l'hôpital.

Le contexte institutionnel et conceptuel du régime du cas psychiatrique dans l'aliénisme

Le médecin classique – Sydenham, pour le plus réputé d'entre eux, ou son modèle antique, Hippocrate - existe entre le corps du livre, abstrait, idéal, d'où il tire son savoir, et le corps du patient, singulier, exposé à lui dans une relation de face-à-face (puisque les malades sont alors des patients privés). Cela vaut pour la médecine dite de l'âme comme pour celle du corps : les ouvrages classiques de Willis, Boissier (*Nosologie méthodique*, 1771)⁵, Cullen (*First lines of the practice of physicks*, 1784⁶) dans leurs pages sur les “ vésanies ” regorgent d'histoires de cas, et on en trouve encore davantage dans le corpus antique. La psychiatrie (qui ne s'appelle pas encore psychiatrie, le terme est de Reil au début du XIX^e siècle, ni même aliénisme, et n'est pas encore distincte de la médecine)

⁵ François Boissier de Sauvages, *Nosologie méthodique, ou distribution des maladies en classes, genres et espèces selon l'esprit de Sydenham et la méthode botanique*, Lyon, Mercier, 1771.

⁶ William Cullen, *First lines of the practice of physicks*, London, Cadell, 1784 (1^{ère} éd. 1777).

est un corpus où les mêmes histoires se transmettent d'ouvrage à ouvrage, indépendamment des modifications doctrinales, venant souvent illustrer des théories différentes. Ainsi, Galien diagnostiquant une passion amoureuse pathologique en prenant le pouls d'un malade est un exemple canonique ; plus tard il y eut Boerhaave guérissant une quinzaine d'orphelins en les épouvantant collectivement. Enfin Willis guérissant George III deviendra l'exemple le plus fameux avant Pinel.

Cabanis, Bichat, Pinel, et les auteurs de la seconde moitié du 18^{ème} siècle en général, expérimentent une autre situation du médecin. En effet, l'horizon sous lequel ils travaillent est avant tout marqué par un événement institutionnel : l'invention de l'hôpital non plus comme lieu d'hospitalité, mais comme lieu de guérison. Les historiens depuis plusieurs décennies ont retracé et réfléchi ce moment majeur de l'histoire médicale, selon des approches souvent divergentes. Toby Gelfand⁷ a magistralement rappelé le fonds social et institutionnel de rivalité entre médecins et chirurgiens et plus généralement entre religieux et médicaux, sur lequel s'édifie cette invention. Akercknecht a rappelé le rôle de l'hôpital dans l'émergence de l'école clinique de Paris⁸. A une échelle plus large, Foucault a insisté sur la manière dont ce qu'on appelle la médicalisation de l'hôpital fut essentielle dans la manière dont les populations sont devenues des enjeux majeurs d'action politique à l'époque moderne ; Scull et Rosen ont montré comment cette mission thérapeutique de l'hôpital s'inscrivait dans les impératifs sociaux neufs liés au capitalisme industriel en formation⁹. Plus encore que Scull et Rosen, Foucault aura donc soutenu, à rebours d'Akercknecht et Gelfand, que la médecine clinique et l'hôpital ne sont pas simplement nés lorsque les médecins, via la chirurgie, se sont reprochés des patients individuels grâce à l'hôpital, au lieu de d'évoluer dans un monde de théories et d'entités nosologiques abstraites. Récemment, Antoine Ermakoff a mis au jour la convergence des logiques proprement médicales, économiques et politiques qui dès la fin du 18^{ème} siècle et dans les

⁷ Toby Gelfand, *Professionalizing Modern Medicine. Paris Surgeons and Medical Science and Institutions in the 18th Century*, Westport, Connecticut and London, Greenwood Press, 1980.

⁸ Erwin Akercknecht, *La médecine hospitalière à Paris : 1794-1848*, Paris, Payot, 1986.

⁹ Foucault, *Histoire de la folie* ; “ La politique de la sante au XVIIIe siècle ”, in: *Dits et écrits, 1954-1988. II, 1976-1988*, Paris, Gallimard, 2001, p. 725–742 ; Foucault, “ L'incorporation de l'hôpital dans la technologie moderne ”, in: *Foucault, Dits et écrits, II*, p. 508–521.

20 premières années du suivant a modelé le fonctionnement et l'idée même de l'hôpital moderne au travers d'une analyse du Conseil Général des Hospices de Paris.¹⁰

C'est donc dans cet hôpital, institution phare en phase d'émergence, qu'on rencontre le cas clinique. A partir des années 1830, pour ce qui est des malades mentaux il s'agit de l'asile psychiatrique, mais Pinel, arrivant à Bicêtre en 1793, puis à la Salpêtrière en 1802, sera responsable à proprement parler d'un département d'aliénés dans un hospice. Il y initiera cette dynamique de revendication d'un hôpital uniquement psychiatrique, que relayeront Esquirol et ses disciples (Leuret, Georget, les Falret, Fodéré, etc.) et qui aboutira à la loi de 1838 stipulant conditions d'internement et ordonnant un établissement psychiatrique par département.¹¹

A l'hôpital, le cas n'est plus cet individu qui incarne un être abstrait, comme le faisait le patient de Sydenham ou de Boissier de Sauvages (*Nosologie méthodique*) pour les espèces ontologiques qu'étaient les "maladies" dans la médecine nosologique : il est plutôt tiré d'une collection de malades, un exemplaire dans une série. La relation exemple / idée qui caractérisait le cas clinique dans les nosologies classiques ne vaut plus pour le cas clinique qu'écrivent les médecins de la fin du XVIII^e siècle, parce qu'il s'agit ici avant tout du rapport d'un individu à une collection : ce rapport est instancié dans un corps malade, et cette collection est, de son côté, constituée selon des procédures précises par l'institution de l'hôpital : relevés réguliers de divers paramètres, enregistrement d'une histoire médicale à l'admission, etc.¹² Le cas, dans ce contexte, signifie un triplet structuré : un nom, un lot de symptômes perceptibles successifs, et un relevé écrit de diverses mesures quotidiennes. Le travail clinique vise à conférer à ce triplet une consistance intelligible. Le cas clinique dans la médecine qui se fait jour en France (au moins) à partir des années 1750 appartient donc à un double espace : espace idéal des espèces pathologiques, dans les nosologies dont celle de Cullen (1769) ou celle de Boissier (1771)

¹⁰ Antoine Ermakoff, *Le conseil général d'administration des hospices civils de Paris : Science d'administration des hôpitaux et médecine clinique (1801-1832)*, thèse, Université Paris 7, Juin 2012.

¹¹ Voir sur ce point Robert Castel, *L'ordre psychiatrique. L'âge d'or de l'aliénisme*, Paris, Minuit, 1976, et Jan Goldstein, *Console and Classify : The French Psychiatric Profession in the Nineteenth Century*, Cambridge, Cambridge University Press, 1987.

¹² Antoine Ermakoff (*ibid.*, sect. III) donne une étude des minutieuses procédures spécifiquement administratives propres à l'hôpital qui contribuent à constituer le "cas clinique" que les ouvrages médicaux relateront ensuite. Ces "technologies de papier" obéissent à des impératifs économiques, utilisent de outils statistiques rudimentaires, et au final définissent une "hospitalisation de la médecine" dont cet auteur montre qu'elle joua un rôle majeur dans l'émergence de la médecine anatomo-clinique.

constituent les meilleurs exemples, les références de l'époque ; et espace social de l'hôpital, où le cas se prélève sur la collection, tel qu'on peut le voir dans la *Médecine clinique* qui redouble la nosologie pinellienne pour en faire une sorte manuel destiné au médecin hospitalier.

Typique de l'articulation de ces deux espaces, est en effet la complémentarité dans l'œuvre de Pinel entre la *Nosographie philosophique* de 1797, et la *Médecine clinique*, censée l'accompagner en présentant de nombreux cas cliniques susceptibles d'aider le nouveau médecin à se reconnaître dans la jungle des espèces de maladies analysées par le premier ouvrage, et en les appliquant à la réalité hospitalière. Ouvrant la *Médecine clinique*, ce recueil censé donner une base en quelque sorte concrète, pratique à la nosographie philosophique, Pinel écrit ainsi : “ Il serait difficile d'exprimer la fluctuation d'opinion, l'incertitude et l'embarras extrême que j'éprouvai, il y a environ douze ans, lorsque je fus appelé à exercer la médecine dans les hospices. Je devais chercher naturellement, dans mes visites ordinaires des malades, à me rendre un compte sévère de tout ce que j'observais, et à éviter du moins les erreurs dangereuses. Mais que d'obstacles passaient en foule sur mes pas par la confusion des objets ! En effet, quel tableau disparate et sans cesse mobile n'offre point un rassemblement de cent cinquante à deux cent malades atteintes de symptômes simultanés ou successifs plus ou moins graves, les uns dus au caractère spécifique des maladies, d'autres aux localités ou à des dispositions individuelles, certains enfin à l'influence spéciale des saisons et de l'atmosphère ! Pouvais-je diriger ma marche d'après des histoires particulières si souvent surchargées de détails superflus ? ”¹³ Le fouillis des sens qu'offre l'accumulation de malades à l'hôpital doit, par le va-et-vient entre la *Nosographie* et la *Médecine clinique*, se laisser organiser dans l'esprit du médecin débutant en une série de cas cliniques reliés entre eux par de relations rationnelles. (Notons que l'histoire de l'institution hospitalière va consister à produire en son dedans un mouvement similaire, par la séparation des malades en territoires pathologiques distincts, la mise en ordre (pour reprendre les termes pinelliens) d' “ histoires particulières ” soutenue par l'atténuation des “ détails superflus ”.¹⁴) Le cas clinique, en médecine pure, et en tant que

¹³ *Médecine clinique*, p. I.

¹⁴ Sur ce point voir George Weisz, *Divide And Conquer : A Comparative History Of Medical Specialization*, Oxford, Oxford University Press, 2006 ; Ermakoff, *op. cit.*, sect. 2.

catégorie discursive, est un vecteur d'ordre dans le fouillis d'individualités qu'est une collection de malades hospitalisés ; mais un ordre intermédiaire entre ce désordre de la première impression et l'ordre idéal de la taxonomie. Le premier vecteur de cette transformation est alors l'enregistrement plus ou moins minutieux des modifications des variables caractérisant l'état d'un individu au long de son hospitalisation, étant entendu que ce relevé n'est qu'un membre d'une large classe de relevés censés correspondre à des malades atteints de pathologies analogues.

Avant d'étudier comment le régime du cas aliéniste diffère de ce cas clinique médical, on doit toutefois noter que le discours médical au sens strict n'est pas homogène, en particulier justement pour ce qui concerne le rapport à l'hôpital. Une citation de l'*Anatomie pathologique* de Bichat établira le contraste avec le cas pinellien : “ On observe l'inflammation dans les gaines des ligaments du carpe et du doigt, etc. Il y a un panaris qui gît évidemment dans la gaine de fléchisseurs de la main, et y produit de grands ravages ; le doigt se gonfle sans rougeur, mais avec une douleur vive qui répond à la face palmaire du doigt. Bientôt du pus s'amasse, et se trouve comprimé entre les os et la synoviale. Il paraît que c'est à cette cause qu'il faut rapporter la plupart des accidents de ce panaris. On trouve, à l'ouverture, toute la capsule rouge et remplie par un pus ichoreux et sanieux. ” (III, XVIII, p. 308, 1804) La médecine anatomo-clinique qu'initie Bichat, on le voit, évite le cas, puisqu'elle porte sur des cadavres (en anatomie pathologique), que le discours relie de manière lâche et *non-individualisée* aux malades qu'ils furent¹⁵. C'est ce type de discours, caractérisable par son investissement fort des qualités sensibles et son insistance sur la précision phénoménologique du récit, que Foucault dans *Naissance de la clinique*¹⁶ prend pour typique du glissement épistémologique introduit par la médecine clinique. Mais comme on le voit sur ces exemples, à l'inverse, la médecine clinique de Pinel, de même que son aliénisme, ne sont pas centrés sur (ou référés en premier lieu à) la salle d'autopsie ; le cas clinique y est prégnant, et relié à l'hôpital où le médecin rencontre ses malades,¹⁷ de

¹⁵ Sur la place de l'anatomie pathologique chez Bichat, voir Philippe Huneman, *Bichat. La vie et la mort*, Paris, PUF, 1998, chap. 1.

¹⁶ Michel Foucault, *Naissance de la clinique*, Paris, Presses universitaires de France, 1963, p. 200 sq.

¹⁷ Sur quelques grandes lignes concernant l'histoire du cas en psychiatrie voir Paul Bercherie, *Les fondements de la clinique, Histoire et structure du savoir psychiatrique*, Paris, Seuil, 1980 ; sur la spécificité de la maladie mentale dans les structures du “ medical care ” au XVIIIème siècle, voir William Bynum, “Health, Disease and Medical Care”, in G.S. Rousseau et R. Porter (eds), *The Ferment of Knowledge*, Cambridge, Cambridge University Press, 1980, p. 211-54.

sorte que l'existence discursive de la maladie n'est ni celle de l'épure phénoménologique des lésions observables indifféremment avant qu'après la mort, ni la tentative de ressaisir une essence idéale de maladie telle que l'entendait la médecine de l'âge classique. Certes, sa *Nosographie* prolonge bien le projet nosologique de cette dernière, mais son articulation avec la *Médecine clinique*, qui en fait dans le même temps un manuel de médecine hospitalière, sort en quelque sorte l'entité pathologique pinellienne de l'espace idéal des taxonomies médicales qui précédèrent. Là où pour Bichat l'ancrage de la pathologie dans la salle d'autopsie induisait le passage de la médecine classique, avec ses espèces, vers la médecine anatomo-clinique, l'articulation avec l'hôpital via la *Médecine clinique* tire la médecine pinellienne hors du cadre nosologique classique d'où semblait s'enraciner sa *Nosographie*.

Médecine clinique, nosographie philosophique, Traité médico philosophique de l'aliénation mentale, chacune de ces œuvres pinelliennes s'élabore à sa manière autour de la mise en ordre de cas. Pourtant le *Traité* ne présente pas le type de mise en ordre que le couple *Médecine clinique/Nosographie* peut opérer. Il y a un écart, souvent constaté¹⁸ entre la pratique de Pinel et la médecine clinique dont il est contemporain et dont sa propre *Nosographie philosophique* a été une référence en matière de méthodologie.¹⁹

De fait, le Pinel du *Traité médico-philosophique*, le Pinel aliéniste, partant de la même expérience hospitalière que la *Médecine clinique*, a un projet différent. Il entend instituer l'aliénisme, c'est-à-dire la psychiatrie comme discipline autonome, pourvue de son objet, de ses modes d'intervention et de son territoire propre. On sait bien qu'il y eut du traitement moral avant Pinel, qu'il s'agisse de la pratique médicale de Tuke à la Retraite, de Haslam à Bedlam, ou même de certaines pages de Boissier de Sauvages, mais le fait qu'il nomme " traitement moral " un ensemble de pratiques qu'il synthétise, ajouté à ceci qu'il attribue à un seul expert – l'aliéniste - la maîtrise de ce traitement, signifie qu'il constitue une relation légitime nouvelle entre un spécialiste médical, d'un côté, et une maladie particulière de l'autre - relation régie par des règles propres. C'est donc à la fois un projet médical, et un projet qui s'écarte justement de ce qui s'institue dans les mêmes

¹⁸ Pigeaud, *Aux portes de la psychiatrie*; Michel Foucault, *Le pouvoir psychiatrique*, Cours au Collège de France (1976), Paris, Seuil, 2003.

¹⁹ On sait que Bichat se réclame de lui pour l'intuition fondatrice de l'anatomie pathologique, à savoir l'attribution des inflammations à des perturbations locales des tissus, et l'identification des unes aux autres.

années du côté de la médecine clinique, dont l'élaboration s'appuie à la fois sur l'anatomie pathologique et sur l'hôpital comme machine à produire et (de plus en plus nettement après 1800) préordonner des collections de cas. Se pencher sur la spécificité du régime du cas dans le *Traité* par rapport à la pratique des grands livres de médecine clinique pinelliens c'est donc demander comment l'aliénisme, au moment même où il fait de la folie, une chose médicale, une maladie mentale, s'institue selon des règles et des procédures qui justement ne relèvent pas de la médecine clinique alors en cours d'établissement. Avant d'examiner ces cas proprement dits, esquissons le second point contextuel qui éclaire leur situation, soit l'univers conceptuel où Pinel, comme Bichat ou Cabanis, évolue.

Contexte conceptuel : projet idéologique et langue neuve

Pinel et Bichat sont des Idéologues²⁰ ; Bichat fut lui-même l'élève de Pinel, Pinel était un proche de Cabanis, le titre exact de sa *Médecine clinique* est “ médecine clinique rendue plus précise et plus exacte par l'application de l'analyse ” ; l'anatomie de Bichat est quant à elle une quête des éléments ultimes – les types de tissus – dont l'assemblage constitue les organes puis les “ appareils ” (= systèmes organiques).

Le mot de l'analyse caractérise l'appartenance commune de ces médecins à l'Idéologie. La tâche scientifique essentielle est l'analyse, c'est-à-dire, aussi bien identifier les éléments épistémologiquement ultimes, que recomposer la réalité concrète à partir de ceux-ci. La science de Bichat comprend ainsi : l'anatomie générale qui identifie les éléments, l'anatomie descriptive qui les recompose pour donner le schéma des différents organes, et enfin la physiologie, qui met à jour, par des procédés expérimentaux décomposant les processus de mort en étapes bien définies, les corrélations interorganiques.²¹ La médecine de Pinel, elle, réactive ce cadre idéologique aussi bien dans le domaine de la psychiatrie, que dans le domaine de la nosologie : la science cherche avant

²⁰ Au sens bien sûr de l'école philosophique de Cabanis, Destutt de Tracy, etc. Sur ce lien à l'Idéologie voir Kathleen Grange, “ Pinel and XVIIIth century psychiatry », *Bulletin of the History of Medicine*, vol. 35, 1961, p. 442-453.

²¹ Huneman, *op.cit.*

tout les espèces simples de maladies, dont d'autres maladies sont des combinaisons complexes.

“ Chacun de ces faits exposé avec précision et avec exactitude, offre ainsi un tableau rapproché propre à être embrassé d'un seul coup d'œil, et à être comparé facilement avec tout autre tableau analogue. Il en est de même de plusieurs de plusieurs faits renfermés sous le titre d'une espèce particulière, de plusieurs espèces réunies en un genre, de plusieurs genres rapportés à un ordre, ou de plusieurs ordres rassemblés sous le titre général d'une classe (...) Les maladies ainsi étudiées dans le rapport de leurs affinité, forment un enchaînement naturel d'idées, sont classifiées d'après leurs signes extérieurs comme tous les autres objets d'histoire naturelle, et finissent par être soumises à des dénominations exactes et invariables. ” (*Médecine clinique*, ix)

Le rôle du récit de cas est alors prescrit par cette méthodologie :

“ Telle est l'origine des histoires multipliées que je donne des maladies aiguës dans cet ouvrage, soit pour servir de fondement à mes principes de nosographie, ou de terme de comparaison pour l'étude de la clinique, soit pour faciliter l'application de l'analyse et de la distinction de ce que j'appelle espèces simples ou espèces compliquées, soit enfin pour faire voir qu'une suite quelconque de maladies bien observées et bien décrites peut être réduite en un ordre aussi régulier et aussi méthodique qu'aucun autre objet d'histoire naturelle. Il faut seulement écarter de leurs notions toute opinion hypothétique, se borner pour leurs signes aux seules impressions faites sur les sens, et considérer chaque maladie comme formant un tout unique résultant de l'ensemble et de la succession de ses symptômes. ” (*ib.*, vii).

Ainsi, l'espèce simple qu'est la péripneumonie, maladie du tissu cellulaire, donne lieu à une observation (*Médecine clinique*, Genre XXIII, ch. 13, p. 131), puis la péripneumonie gastrique donne lieu à une histoire de cas (*ib.*, 135 sq) ; le texte met en regard l'un de l'autre les symptômes afin de donner lieu à une sorte de corrélation générale logique des espèces et des genres selon leurs symptômes. (cf. Fig. 1)

ESPÈCES COMPLIQUÉES.

Péricnephronie gastrique.

Depargi, âgée de cinquante-deux ans, éprouve un frisson avec assoupissement.

2^e jour de la maladie. Au réveil, chaleur, douleur vive sous les côtes asternales droites ; les jours suivans, paroxysme, rougeur des pommettes.

<i>Symptômes péricnephroniques.</i>	<i>Symptômes gastriques.</i>	<i>Symptômes communs.</i>
Gr. Douleur poignante au côté droit de la poitrine, gêne de la respiration, toux, expectoration muqueuse.	Douleur pulsative à la tête ; bouche pâteuse, amère ; langue recouverte d'un enduit jaunâtre ; pesanteur à l'épigastre.	Chaleur forte de la peau, puis pleur, fréquent, paroxysme le soir. L'émétique provoque des évacuations abondantes.

(Fig.1)

Les narrations de cas ont alors une autre fonction, de par la précision de leur situation et de leur contexte : elles montrent l'hôpital lui-même comme responsable de la variété dans les cas :

“ Les maladies dont je rapporte l'histoire, comparées avec les mêmes espèces observées dans d'autres hôpitaux, peuvent offrir des différences frappantes, soit pour l'intensité, soit pour la réunion ou les modifications de certains symptômes qui peuvent tenir de part et d'autre à l'influence des localités. Il importe donc d'indiquer ici les principaux traits de la position topographique de l'hospice, de sa distribution intérieure, du nombre, de l'état particulier et de la manière de vivre des infirmes qui sont venues y chercher un asile. ”
(*Médecine clinique*, x).

On voit donc comment l'institution hospitalière, de manière générale, est toujours présente dans la constitution même du cas clinique ; dans la suite nous monterons comment cette présence prend une tonalité spécifique dans le cas proprement psychiatrique.

Parallèlement au médecin clinique, l'aliéniste poursuit une tâche scientifique par son investissement dans la tâche d'analyse. Certes – et cela se précisera dans la suite – les unités d'analyse ne sont pas aussi facilement spécifiées que celles du discours médical, puisqu'il ne suffit pas de l'acuité perceptive pour embrasser l'ensemble des traits qui définissent une catégorie pathologique, n'importe quel trait comportemental ou cognitif pouvant être partie intégrante d'une maladie mentale. Néanmoins le projet aliéniste même participe au même effort d'analyse et de reconstitution des espèces de maladie.

Le régime pinellien du cas clinique

2.1. Généralités

Sur le fond de ce projet intellectuel, il s'agit donc de déterminer l'usage et la grammaire des récits de cas du *Traité*. Pour commencer donnons deux exemples de ces cas, dont je tirerai quelques points communs très généraux. Ensuite j'approfondirai deux dimensions essentielles, le temps et la causalité d'une part, la place de l'hôpital de l'autre. Je convoquerai au fil de l'analyse plusieurs cas qui illustrent chacun de ces points particuliers. Le dernier volet de l'étude abordera la fonction rhétorique du cas pinellien.

Cas A. “ Une mobilité turbulente et incoercible, une succession rapide et comme instantanée d'idées qui semblent naître et comme pulluler dans l'entendement, sans aucune impression faite sur les sens ; un flux et reflux continu et ridicule d'objets chimériques qui se choquent, s'alternent et se détruisent les uns les autres sans aucune intermission et sans aucun rapport entre eux ; le même concours tumultueux d'affections morales, de sentiments de joie, de tristesse, de colère, qui naissent fortuitement et disparaissent de même, sans laisser aucune trace et sans avoir aucune correspondance avec les impressions des objets externes ; tel est le caractère fondamental de la démence dont je parle. Un homme doué d'un patriotisme ardent, mais peu éclairé, et qui était un des plus zélés admirateurs du fameux Danton, se trouve présent à la séance du Corps législatif où fut prononcé le décret d'accusation contre ce député ; il se retire dans une sorte de consternation et de désespoir, reste renfermé chez lui plusieurs jours, livré aux idées les plus sinistres et les plus mélancoliques. “ Comment ? Danton, un traître ! répète-t-il sans cesse ; on ne peut plus se fier à personne et la République est perdue ” ! Plus d'appétit, plus de sommeil, et bientôt l'aliénation la plus complète ; il subit le traitement usité au ci-devant Hôtel-Dieu, et il est conduit à Bicêtre. Je l'ai gardé plusieurs mois aux infirmeries de cet hospice, livré à une sorte de rêvasserie douce, à un babil confus et ininterrompu de termes les plus disparates ; il parlait tour à tour de poignards, de sabres, de vaisseaux démâtés, de vertes prairies, de sa femme, de son château, etc. ; il ne songeait à manger que lorsqu'on mettait ses aliments dans sa bouche, et il était absolument réduit à une existence automatique. ” (TMP, II, §XVIII p. 162)

Cas B. “ Un jeune homme, consterné du renversement du culte catholique en France, et dominé par des préjugés religieux, devient maniaque, et après le traitement usité de l'Hotel-Dieu, il est transféré à Bicêtre. Rien n'égale sa

sombre misanthropie ; il ne parle que des tourments de l'autre vie, et il pense que pour s'y soustraire, il doit imiter les abstinences et les macérations des anciens anachorètes ; il s'interdit dès lors toute nourriture, et vers le quatrième jour de cette résolution inébranlable, son état de langueur fait craindre pour sa vie ; remontrances amicales, invitations pressantes, tout est vain ; il repousse avec dureté un potage qu'on lui sert, et il affecte d'écarter la paille de sa couche, pour reposer sur les planches. Le cours de ses idées sinistres pouvait-il être autrement détruit ou contrebalancé, que par l'impression d'une crainte vive et profonde ? C'est dans cette vue que le citoyen Pussin se présente le soir à la porte de sa loge, avec un appareil propre à effrayer, l'œil en feu, un ton de voix foudroyant, un groupe de gens de service pressés autour et armés de fortes chaînes qu'ils agitent avec fracas ; on met un potage auprès de l'aliéné, et on lui intime l'ordre le plus précis de le prendre pendant la nuit, s'il ne veut pas encourir les traitements les plus cruels ; on se retire, et on le laisse dans l'état le plus pénible de fluctuation, entre l'idée de la punition qui le menace, et la perspective effrayante des tourments de l'autre vie. Après un combat intérieur de plusieurs heures, la première idée l'emporte, et il se détermine à prendre sa nourriture. On le soumet ensuite à un régime propre à le restaurer ; le sommeil et les forces reviennent par degrés, ainsi que l'usage de la raison, et il échappe de cette manière à une mort certaine. C'est durant sa convalescence qu'il m'a fait souvent l'aveu de ses agitations cruelles et de ses perplexités durant la nuit de son épreuve. ” (TMP, II, VIII, p. 59 sq)

(a) Le cas clinique, dans l'aliénisme, a d'abord cette propriété d'exemplifier par des comportements et des croyances précises une espèce d'aliénation : la démence, en A, est une anomalie dans la production d'idées ; le cas cité illustre cette anomalie par des idées septiques.

(b) Ensuite, le cas en quelque sorte historicise cette espèce, il la déploie dans le temps : à une forme de dérèglement idéal, il fait correspondre une suite d'épisodes comportementaux et imaginatifs. Comme dans le paradigme hippocratique (prodromes, maladie qui se déclare, crise, résolution), le cas marque au cours de cette séquence des crises et des évolutions, même si dans ces exemples manque l'idée de “ jour critique ”, c'est-à-dire de rythme prédéterminé, propre à la maladie. Pour ce qui est de la fin de la maladie, elle peut clore le cas, ou bien ne pas intervenir de sorte que la maladie reste en suspens : ainsi, en B, notre malade guérit par un subterfuge asilaire.

(c). Enfin, c'est d'être intégré dans le cas qu'une idée soutenue par un patient, qu'un comportement démontré, que tout fait subjectif devient pathologique : à cette condition,

n'importe quel trait – croyance, actes, dérèglements organiques – peut faire symptôme. Le cas se construit ainsi par accumulation de faits de tous ordres (parler de poignards, ne pas manger, etc.) qui deviennent symptomatiques par leur inclusion dans ce déploiement temporel. Tout est clinique dans le cas psychiatrique, et ceci illustre la généralité de la maladie mentale, une des thèses majeures de l'aliénisme : ni physique ni organique, elle affecte l'homme "au physique et au moral", pour reprendre la formule d'un autre Idéologue.²²

A partir de là, on peut donc repérer quelques traits de la grammaire et de la fonction du cas clinique ; tous ne les partagent pas, mais ils sont chacun constitutifs de ce régime pinellien du cas dont à la fin de l'étude je tenterai de montrer la solidarité avec le projet aliéniste même.

2.2. Le temps et la causalité.

Le cas est une histoire, donc une séquence temporelle ; quelles en sont les règles d'énonciation ? Les cas d'aliénation ont toujours un début ; celui-ci peut être une cause identifiée, ramassée parfois en l'espace d'un événement, d'une parole ; il peut aussi avoir lieu comme un long processus, dont le rapport causal avec la folie elle-même se résume à en être l'antécédent. Je cite ici deux cas exemplaires de ces deux types, dont je reparlerai, et que je nomme l'Ami, et le Vigneron crédule.

Cas C. L'Ami. " *Histoire d'une manie où le traitement moral aurait été nécessaire.* Un jeune homme, âgé de Vingt-quatre ans, et doué d'une imagination ardente, vient à Paris pour suivre le cours de ses études, et se croit destiné par la nature à jouer dans la suite le rôle le plus brillant dans le barreau. Application continuelle, vie passée dans la retraite, sobriété extrême pour donner plus d'essor à ses facultés morales, régime pythagorique, adopté dans toute la rigueur du terme. **Quelques mois après**, migraines violentes, saignements fréquents du nez, resserrements spasmodiques de la poitrine, douleurs vagues des intestins, flactuosités incommodes, sensibilité morale très-exaltée. Quelquefois, il m'aborde avec un air rayonnant de joie, et il ne peut exprimer la félicité suprême qu'il dit éprouver en lui-même : d'autres fois je le

²² Sur l'indétermination physique/organique de Pinel, voir Philippe Huneman " 'Animal Economy' : Anthropology and the Rise of Psychiatry from the *Encyclopédie* to the Alienists ", in L. Wolff, M. Cipolloni eds., *Anthropology of the Enlightenment*, Stanford, Stanford University Press, 2007, p. 262-276, et Swain, *Sujet de la folie*, op.cit.

trouve plongé dans les horreurs de la consternation et du désespoir, et il me fait les instances les plus vives de mettre fin à ses souffrances ; les caractères de l'hypocondrie la plus profonde étaient aisés à reconnaître ; je lui en retrace les dangers pour la suite, et je lui conjure souvent de changer sa manière de vivre ; mais il poursuit toujours son plan avec l'obstination la plus inflexible : augmentation des symptômes nerveux de la tête, du bas-ventre, de la poitrine ; alternatives plus fréquentes d'un abattement extrême et d'une joie convulsive, terreurs pusillanimes, surtout dans les ombres de la nuit, angoisses inexprimables. Il venait quelquefois me trouver, fondant en larmes, et me conjurant de l'arracher des bras de la mort ; je l'entraînais alors dans la campagne, et quelques tours de promenade, avec des propos consolants, semblaient lui rendre une nouvelle vie ; mais à son retour dans la chambre, nouvelles perplexités, terreurs pusillanimes renaissantes ; il trouve un surcroît de désolation et de désespoir dans la confusion croissante de ses idées, l'impossibilité de se livrer désormais à l'étude, et la conviction accablante de voir s'évanouir pour l'avenir la perspective de célébrité et de gloire dont son imagination avait été bercée ; l'aliénation la plus complète suit de près. Un jour qu'il se rend au spectacle pour se distraire, on joue la pièce du *Philosophe sans le savoir*, et dès lors le voilà assailli des soupçons les plus noirs et les plus ombrageux ; il est profondément persuadé qu'on a joué ses ridicules : il m'accuse d'avoir fourni moi-même les matériaux de la pièce et, dès le lendemain matin, il vient me faire les reproches les plus sérieux et les plus amers, d'avoir trahi les droits de l'amitié, et de l'avoir exposé à la dérision publique. Son délire n'a point de bornes ; il croit voir dans les promenades publiques des comédiens travestis en moines et en prêtres, pour étudier tous les gestes, et surprendre le secret de ses pensées. Dans l'ombre de la nuit, il se croit assailli, tantôt par des espions, tantôt par des voleurs et des assassins ; et une fois, il répand l'alarme dans le quartier, en ouvrant brusquement les croisées, et en criant de toutes ses forces qu'on en voulait à sa vie. Un de ses parents le détermine à lui faire subir le traitement de la manie au ci-devant Hôtel-Dieu, et il le fait partir vingt jours après avec un compagnon de voyage pour se rendre dans une petite ville voisine des Pyrénées. Egaleme nt affaibli au moral et au physique, toujours dans les alternatives de quelques écarts du délire le plus extravagant et des accès de sa noire et profonde mélancolie, il se condamne à un isolement profond dans sa maison paternelle : ennui, dégoût, insurmontable de la vie, refus de toute nourriture, brusquerie contre tout ce qui l'avoisine ; il trompe enfin la surveillance de sa garde, fuit en chemise dans un bois voisin, s'égare, expire de faiblesse et d'inanition, et deux jours après, on le retrouve mort, tenant dans sa main le fameux livre de Platon sur l'immortalité de l'âme. " (*Traité*, II, §V, p. 54 sq)

Cas D. Le vigneron crédule.

“ *Exemple d’une mélancolie avec bigoterie.* Un missionnaire, par ses fougueuses déclamations et l’image des tourments de l’autre vie, épouvante tellement un vigneron crédule que celui-ci se croit franchement dévolu aux brasiers éternels, et qu’il ne pense plus qu’à sauver sa famille et à la faire jouir des palmes du martyre, dont une fréquente lecture de la vie des saints lui avait fait les peintures les plus séduisantes. Il essaie d’abord de commettre ce crime horrible sur sa femme, qui parvient à s’échapper de ses mains et bientôt après son bras forcené se porte sur deux enfants en bas âge, et il a la barbarie de les immoler de sang froid, pour leur procurer la vie éternelle. Il est cité devant les tribunaux et, durant l’instruction de son procès, il égorge encore un criminel qui était avec lui dans les cachots, toujours dans la vue de faire une œuvre expiatoire. Son aliénation étant constatée on le condamne à être enfermé pour la vie dans les loges de Bicêtre. L’isolement d’une longue détention, toujours propre à enflammer l’imagination, l’idée d’avoir échappé à la mort malgré l’arrêt qu’il suppose avoir été prononcé par les juges, aggravent son délire, et lui font penser qu’il est *la quatrième personne de la Trinité*, que sa mission spéciale est de sauver le monde par le baptême de sang, et que tous les potentats de la terre réunis ne sauraient attenter à sa vie. Son égarement se borne d’ailleurs à tout ce qui se rapporte à la religion ; car sur tout autre objet il paraît jouir de la raison la plus saine. Plus de dix années se sont passées dans une étroite réclusion, et les apparences soutenues d’un état calme et tranquille, déterminèrent à lui accorder la liberté dans les cours de l’hospice avec les autres convalescents. Quatre nouvelles années d’épreuve semblaient rassurer, lorsqu’on a vu tout-à-coup se reproduire ses idées sanguinaires comme un objet du culte religieux. Cette année, le 10 Nivôse, qui correspond à la veille de Noël, vieux style, il forme le projet atroce de faire un sacrifice expiatoire de tous les hommes de l’hospice ; il se procure un tranchet de cordonnier, saisit le moment où le cordonnier descend pour faire sa ronde, lui porte un coup par derrière, qui glisse heureusement sur les côtes, coupe la gorge à deux aliénés qui étaient à ses côtés ; et il aurait ainsi poursuivi le cours de son homicide, si les gens de service ne fussent promptement venus en forcer pour arrêter sa rage. On n’a pas besoin de remarquer que sa réclusion sera à jamais irrévocable. ” (*Traité*, II, §XV, p. 72 sq)

2.2.1. L'origine et la durée.

Dans les deux cas, tout fait signe, dans ces processus ; de la sorte tout, rétrospectivement, se donne comme les prodromes de la folie. Cette généralité de la folie, se traduit stylistiquement par la concaténation des attributs mis bout à bout, identique à une énumération clinique de symptômes. Ainsi, pour le cas de l'ami, des phrases entières sont faites de groupes nominaux séparés par des virgules. On retrouve la même forme grammaticale dans les cas de maladies dans la *Médecine clinique* :

“ Depuis quelques jours, perte de l'appétit, dégoût, malaise. Dans la nuit du 11 au 12 germinal, céphalalgie très violente, frissons, douleurs abdominales intolérables, nausées ; enfin, vomissements avec secousses vives, déjections alvines abondantes. Les matières rendues sont vertes, et celles qui s'échappent de l'anus sont, en outre, mêlées de sang. Le lendemain à huit heures, le vomissement cesse, le dévoiement continue, mais les matières ne sont plus sanguinolentes. Les douleurs de colique persistent encore ; enduit jaune de la langue, peau sèche ; le soir, léger frisson, suivi d'un peu de chaleur avec moiteur. ” (*Médecine clinique*, p. 23, cas de la fièvre méningogatsrique.)

La langue du cas psychiatrique est donc la langue même du cas de la médecine clinique : énumération de signes, sans liens de subordination. Ces signes, notons-le, sont indifféremment moraux ou physiques : en effet, selon la conception même de Pinel, la folie est une “ perturbation générale de l'économie animale ”.²³ Plus exactement, elle n'est ni spécialement organique – Pinel montrera qu'on ne peut pas la relier nécessairement à des lésions du cerveau – ni spécialement mentale : puisqu'elle concerne l'économie animale dans son ensemble, elle se passe rarement d'atteintes au corps, et peut être causée et soignée par des moyens en partie physiques²⁴... Mais le “ *quelques mois après*, migraines violentes... ” indique en même temps un lien causal sous-entendu, que le lecteur est à même de saisir alors qu'aucun terme de la phrase ne le représente : Pinel fait ici appel à cette

²³ Cette expression est reprise et analysée dans Philippe Huneman, “Montpellier Vitalism and the Emergence of Alienism in France (1750—1800) : The Case of the Passions”, *Science in Context*, 2008, vol. 21, n° 4, p. 615-647. Voir aussi Swain, *op.cit.* et Swain et Gauchet, *op.cit.*

²⁴ Sur cette notion d' “ économie animale ” voir Huneman, “ Montpellier vitalism... ”, et “ ‘Animal economy’... ”, ainsi que Charles Wolfe et Terada (2008), “The animal economy as object and program in Montpellier vitalism”, *Science in Context*, vol. 21, n° 4, p. 537-79 ; sur son rôle dans la naissance de l'aliénisme français, Huneman “ ‘Animal economy’.. » et “ Les théories de l'économie animale et la naissance de l'aliénisme ”, *Psychiatrie, Sciences humaines, Neurosciences*, II, 2, 2004, p. 47-59.

vieille thématique, la folie induite par l'excès d'études, un lieu commun de la médecine des fous depuis Arétée.²⁵

Toutefois, l'événement causateur ne fait pas totalement défaut dans le cas de l'Ami, et constate ici une parenté avec le cas du Vigneron crédule, rendu fou apparemment par le discours d'un missionnaire. Pinel écrit en effet de celui-là : “ Un jour qu'il se rend au spectacle pour se distraire, etc. ”. Ainsi, si dans ce cas, le départ d'avec la normalité – les petits désagréments de l'étude – se fait progressivement, et au fond de manière incompréhensible, il faut toutefois, pour rendre palpable l'entrée dans la folie, un événement auquel on attribue le passage à la folie, la discontinuité qui sépare le normal de la maladie. Après le spectacle, écrit donc Pinel, “ son délire n'a plus de bornes ”. Bien sûr, pourrait-on dire, la cause ici est reconstruite, la pièce de théâtre n'a été qu'un pas de plus dans un long processus, et sans doute, autre chose, sinon, aurait précipité la cure. Mais l'essentiel est qu'il faille un événement, un déclencheur, pour séparer la normalité de la folie, en même temps qu'indiquer une origine du mal. Ceci par contrecoup, donne à la folie un aspect arbitraire : elle est comme tendue entre le hasard de l'événement déclencheur et le mystère du processus de détérioration qui l'a précédé et dont rien ne vient expliquer les raisons.

Dans de nombreux cas, les événements originaux – le discours sur Danton, en A ; le discours séduisant le vigneron crédule – n'ont pas de vertu explicative, au sens nomothétique, rien en eux ne porte nécessairement la folie ; au contraire, ils sont même banals. A l'inverse, le fait qu'ils puissent avoir de telles conséquences indique déjà la folie. En somme, ils provoquent moins une folie qu'ils ne la manifestent, puisque par un truisme clair, il faut être déjà fou pour être rendu fou par une pièce de théâtre. La catégorie fondamentale de l'entrée en folie, c'est donc la *révélation*. La plupart de ces cas seront donc construits selon une logique du caché/révéle. Certes, en médecine, le cas clinique est toujours apparition de quelque chose d'auparavant enfoui dans la profondeur des corps ; mais le cas psychiatrique met en oeuvre cette logique selon des paramètres qui lui sont propres.

Si de manière générale, l'événement original n'explique rien, le propos de Pinel

²⁵ Et que Tissot avait remis au goût du jour avec son célèbre *De la santé des gens de lettres* (1768) ; merci à Patrick Singy pour cette référence.

n'est pas de comprendre la cause de la folie. Celle-ci reste mystérieuse, elle est reconstruite en tant qu'origine par le médecin, non pas à partir d'éléments théoriques qu'il aurait auparavant élaborés, mais à partir d'informations factuelles et d'éléments de sens commun unanimement admis. Il s'agit souvent ici d'une sorte de compréhension narrative commune à tous, au sens où l'on comprend des choses comme " la déception sentimentale l'a rendu fou ", etc. Pour cette raison, certains cas peuvent être complets, mentionnant la cause ; d'autres ne le sont pas ; d'autres ne donnent qu'un fragment de la folie, pour illustrer une espèce de maladie mentale.

L'origine de la folie est ainsi reconstruite ; il s'ensuit une narration dont la temporalité est scandée par les crises, une temporalité parfois longue, à la différence des cas de la médecine clinique.²⁶ Cette longue durée est aussi visible dans le cas du vigneron, elle est scandée par les changements de pôle du délire, ou bien par le passage du latent au manifeste : " Plus de dix années se sont passées dans une étroite réclusion, et les apparences soutenues d'un état calme et tranquille. (...) quatre nouvelles années (...) ". Dans ce cas, il est notable que le fou, sur le long terme, semble guéri : traditionnellement, la pénitence est une pratique religieuse, et dans cette perspective l'isolement rachète, fait rédemption, et ici, il n'est évidemment pas indifférent que le fou soit un maniaque religieux. Ainsi, selon cette logique chrétienne, et selon les " apparences ", le fou devrait être sauvé, guéri. Or il n'en est rien : " on a vu tout à-coup {et l'adverbe, ici, s'oppose bien sûr aux " dix années, etc. "} se reproduire ses idées sanguinaires comme un objet du culte religieux ". Construire un cas, c'est donc articuler d'une part, le latent et le manifeste, de l'autre, la longue durée (dix ans, etc.) et le " soudain ", le " tout à-coup " ; ce lexique de l'instantané désigne le

²⁶ " Un aliéné de cette espèce (mélancolie), confié à mes soins depuis douze années, et déjà avancé en âge, a déliré pendant les huit premières que sur l'idée chimérique d'un prétendu empoisonnement dont il se croyait menacé. Dans cet intervalle de temps, nul écart de sa conduite, nulle autre marque d'aliénation ; il était même d'une réserve extrême dans ses propos, persuadé que ses parents cherchaient à le faire interdire et à s'emparer de ses biens ; l'idée d'un prétendu poison le rendait seulement très ombrageux, et il n'osait manger que les aliments pris à la dérobée dans la cuisine de son pensionnat. Vers la huitième année de la réclusion, son délire primitif a changé de caractère. Il a cru d'abord être devenu le plus grand potentat du monde, puis l'égal du créateur et le souverain du monde. Cette idée fait encore sa félicité suprême. " (*Traité*, p. 146).

plan du surgissement du latent, de la révélation. Mais bien entendu, et on y reviendra, cette longue durée suppose un appareil continu de contrôle du fou, elle suppose qu'il soit interné.

2.2.2. *Conflits*

L'opposition de la longue durée et de l'instant rejoint celle de l'enfoui et du manifeste dans une situation très précise, qui revient souvent dans le traité, et qui est caractéristique du cas clinique : le motif du *combat intérieur*.

Considérons quelques cas : d'abord le "combat intérieur" du cas B, l'homme qui ne veut plus manger, devant l'alternative où le surveillant l'a placé.

Soit ensuite le cas d'un homme "livré autrefois à un art mécanique", qui se trouve "renfermé maintenant à Bicêtre", cas exemplaire de manie sans délire, individu parfois traversé d'accès de fureur, durant lesquels : "s'il peut saisir un instrument tranchant, il est porté à sacrifier avec une espèce de fureur la première personne qui s'offre à sa vue. Il jouit cependant à d'autres égards du libre exercice de sa raison, même durant ses accès ; il répond directement aux questions qu'on lui fait et ne laisse échapper aucune incohérence dans les idées, aucun signe de délire" (*Traité*, IV, §X, 152). Cette ambivalence se traduit par le combat intérieur : "à Bicêtre, même accès de fureur périodique, mêmes penchants automatiques à des actes d'atrocités, dirigés quelquefois contre le surveillant, dont il ne cesse de louer les soins compatissants et la douceur. Ce combat intérieur que lui fait éprouver une raison saine en opposition avec une cruauté sanguinaire le réduisent quelque fois au désespoir" (*ib.*, 153).

Le combat intérieur est donc aussi bien (en B) le propre de la manie – avec une issue toujours aléatoire – que ce qu'induit le médecin pour guérir le maniaque. En effet guérir, c'est souvent produire par des moyens artificiels ce conflit interne qui accompagne la folie, mais le produire de tel sorte qu'il soit comme biaisé – ainsi la médecine arrive parfois à contrebalancer l'idée délirante par un appui sur la puissance de l'instinct de survie. Ainsi, l'aliéné qui repousse toute nourriture, mais s'excite, transpire et boit, se voit supprimer par le surveillant, "alors qu'il ne conserve plus qu'une apparence de squelette", sa ration d'eau, et substituer "un bouillon gras".

"L'aliéné reste alors flottant entre deux impulsions contraires, l'une est celle d'une soif dévorante qui le porte à avaler un liquide quelconque, l'autre est la

résolution ferme et immuable d'accélérer le terme de sa vie. La première enfin l'emporte, il prend avec avidité le bouillon, et aussitôt il obtient, à titre de récompense, l'usage libre de l'eau froide. ” (*ib.* V, §III, 183)

Bien entendu, ce type de guérison présuppose tout un appareil disponible en arrière-plan pour simuler en quelque sorte des conditions extrêmes sans les réaliser, pour construire, mettre en scène ce conflit que l'aliéné doit intérioriser. Le cas raconte ce conflit et son issue : il est donc impensable sans un tel appareil en quelque sorte théâtral. Il ne s'agit ni de médecines, ni de paroles consolatrices, alternative dans laquelle se situait l'ancienne psychiatrie, usant tour à tour de l'un et de l'autre, de médications pour les folies à origine plus physiques, de réconfort pour les folies comme passions ou issues de passions.²⁷

2.2.3. *Artifices et thérapie.*

Le cas, ici, fournit une preuve de la vision pénétrante de l'aliéniste : lui seul sait que derrière la décision d'un aliéné, il y a joute, il y a possibilité de conflit, et sait comment induire et gérer un tel conflit. On touche du doigt, ici, la sophistication de la pratique bientôt asilaire : le conflit *intérieur*, intime, est paradoxalement induit, contrôlé, fabriqué ; la réalité à laquelle le fou est acculé (faim, soif, danger, etc.) est paradoxale car artificielle, montée de toutes pièces. On conçoit donc qu'il faille, pour guérir les fous, des dispositifs spéciaux, qui ne sont pas les admonestations ou drogues de l'ancienne thérapeutique mentale, que définissait cette alternative de la chimie et de la parole. Le cas clinique exige ainsi une institution de type asilaire, dans le même temps qu'il fait la preuve de l'intimité du savoir de l'aliéniste, de sa prise sur l'intime du malade. Tel est bien le sens de “ l'aveu ” de la dernière phrase du cas B : “ qu'il m'a fait souvent l'aveu de ses agitations cruelles et de ses perplexités durant la nuit de son épreuve ” - à Pinel et à nul autre car seul le psychiatre était à même d'entendre le vécu intérieur de ce sujet, confronté à une expérience terrifiante dont à son insu son confident avait été l'artisan.

Pareil cas, dans le *Traité*, suivait celui de l'ami. Or ce dernier était précisément destiné à montrer la nécessité d'un certain traitement, comme l'indiquait son sous-titre : *Histoire d'une manie où le traitement moral aurait été nécessaire*. Ainsi, le malade

²⁷ Sur cette division de la psychiatrie classique, voir Pigeaud, *op.cit.*, ch.1 ; Huneman, “ Passions ”, *art. cit.*

“ trompe la surveillance de sa garde ” pour aller se suicider, ce qui suggère entre autres que le traitement du fou exige d’avoir les moyens de le surveiller en permanence. Le paragraphe qui succède immédiatement commence donc par :

“ Dans le traitement de sa manie, il était en mon pouvoir d’user d’un grand nombre de remèdes ; mais le plus puissant de tous me manquait, celui qu’on ne peut guère trouver que dans un hospice bien ordonné, celui qui consiste dans l’art de subjuguier et de dompter, pour ainsi dire, l’aliéné, en le mettant dans l’étroite dépendance d’un homme qui, par ses qualités physiques et morales, soit propre à exercer sur lui un empire irrésistible et à changer la chaîne vicieuse de ses idées ” (*Traité*, p. 57).

Passage exemplaire, bien sûr : pour guérir le fou, il faut le doublet “ hospice bien ordonné ” – “ étroite dépendance d’un homme”. L’un ne va pas sans l’autre : cette “ emprise irrésistible ”, malgré toutes ses qualités, Pinel n’aura pu l’exercer sur son ami car il n’y avait alors pas l’asile, machine à substituer ses propres chaînes vertueuses à la chaîne vicieuse des idées de l’aliéné. Les cas suivants, dont B, illustrent donc comment cet hôpital et cet homme singulier peuvent casser une telle chaîne. Notons que la référence à “ l’homme ” qui exerce “ l’empire ” sur le fou n’est accompagnée d’aucune mention de ses qualités médicales, il est seulement question de qualités personnelles : un rapport personnel qui soit en même temps interne à un appareil, telle est la définition de la pratique aliéniste telle que l’exige Pinel, et que le cas de l’ami illustre par défaut.

Le cas de l’ami et le cas du vigneron cités en commençant cette section sont des cas paradigmatiques dans le texte car ils sont en quelque sorte les cas-limites. D’abord, ils sont complets, par opposition aux cas incomplets qui omettent soit l’origine soit un épisode soit le dénouement de la pathologie ; ensuite, ils se caractérisent tous deux par leur rapport à l’hôpital. Le premier, parce qu’il aurait été guéri en hôpital, en montre donc l’utilité, la nécessité ; le second, parce qu’aucune médecine ne saurait le guérir et qu’il contraint l’hôpital à assumer sa seconde fonction, à savoir l’internement, non pas qu’il soit l’équivalent d’une prison comme lieu punitif, mais parce que seul l’hôpital, en tant que le médecin s’y trouve, peut prévenir certains individus, apparemment normaux mais en profondeur fous, de provoquer des dommages.²⁸ A l’exemple de ces deux-là tous les cas de

²⁸ La nécessité d’un médecin pour écrire les “ lettres de santé » qui doivent libérer les fous, que Pinel évoque

l'aliéniste tournent donc autour de l'hôpital. Abordons donc maintenant l'hôpital comme paramètre essentiel du régime pinellien du cas clinique.

2.3. Hôpital et guérison.

2.3.1. L'hôpital comme norme.

On doit d'abord mieux situer, institutionnellement, le *Traité médico-philosophique sur l'aliénation mentale ou la manie*. Pinel a été nommé à Bicêtre ; les asiles psychiatriques ne sont pas encore créés, il y a simplement ici un lieu qui contient des fous ; Pinel y expérimente ce qu'il va nommer le "traitement moral" en découvrant sa pratique par Pussin, le surveillant.²⁹ Le *Traité* relate cette pratique, avance une nouvelle définition-classification de l'aliénation et quelques positions théoriques sur sa nature, et revendique un certain type d'institution propre à soigner les fous. Nommé peu après à la Salpêtrière Pinel pourra d'ailleurs entamer la constitution de ce genre d'institution, qui ne sera vraiment une réalité qu'avec la loi de 1838.³⁰ Le *Traité* est donc militant, à la différence de la *Médecine clinique* qui tente, on l'a vu, de reconstruire des espèces de maladie dans le fouillis de cas offerts à la vue du médecin entrant dans l'hôpital. C'est pourquoi le cas de l'ami s'avérait exemplaire : il est présenté comme une faute de l'institution médicale, faute originelle de la médecine de l'âme, une faute à réparer par la réforme proposée par Pinel, dans le sillage du mouvement qui a vu émerger les *Instructions* de Colombier et Doublet en 1785,³¹ les remarques de Tenon dans les *Rapports sur les hôpitaux*³² sur le caractère guérisseur d'une institution pour les fous³³. Mais de manière générale, les cas cliniques du

(*Traité*, p. 292) relève bien entendu de cette logique du "caché ». Sur la capacité du seul aliéniste à discerner derrière les apparences normales qui est guéri de qui ne l'est pas, cf. Huneman, "Les théories de l'économie animale et la naissance de l'aliénisme", p. 56-59.

²⁹ Parmi la masse d'analyses historiques de cette entrée de Pinel à Bicêtre, voir Jacques Postel, *Genèse de la psychiatrie – les premiers écrits de Philippe Pinel*, Paris, Le Sycomore, 1981 ; Jean Garrabé ed. *Philippe Pinel*, Paris, Payot, 1994 ; Weiner, *art. cit.*

³⁰ Castel, *op.cit.*

³¹ Jean Colombier, François Doublet, "Observations faites dans le département des hôpitaux civils", *Journal de médecine, chirurgie, pharmacie*, vol. 63, 1785, p. 1-46, 289-359, 445-483.

³² Jacques Tenon, *Mémoires sur les hôpitaux de Paris*, Paris, chez Royez, libraire, 1788.

³³ Sur ce mouvement de professionnalisation de la psychiatrie et d'institution des hôpitaux, de manière plus générale, voir Swain et Gauchet, *op.cit.*, ch. 4, Bynum, *art. cit.* ; Goldstein, *op.cit.* ; et les analyses sociologiques qui présentent une interprétation en termes de rapports sociaux développées par Andrew Scull, *The Most Solitary of Afflictions : Madness and Society in Britain, 1700-1900*, Haven and London, Yale University Press, 1993 ; "Madness and Segregative Control : The Rise of the Insane Asylum", *Social Problems*, vol. 24, n° 3, 1977, p. 337-351 ; "From madness to mental illness : medical men as moral

Traité établissent toujours des rapports entre la maladie et l'hôpital, et entre l'hôpital et la société. Contrairement aux cas de la *Médecine clinique*, dont le cadre est d'entrée de jeu l'hôpital et dont la relation est nécessairement située dans un tel cadre, le cas selon l'aliéniste a un rapport spécifique complexe avec l'hôpital. Soit un premier exemple :

Cas E. “ *Exemple d'une sorte d'emportement maniaque sans délire.* Une éducation nulle ou mal dirigée, ou bien un naturel pervers et indisciplinable peuvent produire les premières nuances de cette espèce d'aliénation, comme l'apprend l'histoire suivante. Un fils unique, élevé sous les yeux d'une mère faible et indulgente, prend l'habitude de se livrer à tous ses caprices, à tous les mouvements d'un cœur fougueux et désordonné ; l'impétuosité de ses penchants augmente et se fortifie par les progrès de l'âge, et l'argent qu'on lui prodigue semble lever tout obstacle à ses volontés suprêmes. Veut-on lui résister ? Son humeur s'exaspère ; il attaque avec audace, cherche à régner par la force, et il vit continuellement dans les querelles et les rixes. Un animal quelconque, un chien, un mouton, un cheval, lui donnent du dépit, il les met soudain à mort. Est-il de quelque assemblée ou de quelque fête, il s'emporte, donne et reçoit des coups, et sort ensanglanté ; d'un autre côté, plein de raison quand il est calme et possesseur dans l'âge adulte d'un grand domaine, il le régit avec un sens droit, remplit les autres devoirs de la société, et se fait connaître même par ses actes de bienfaisance envers les infortunés. Des blessures, des procès, des amendes pécuniaires avaient été le seul fruit de son malheureux penchant aux rixes ; mais un fait notoire met un terme à ses actes de violence : il s'emporte un jour contre une femme qui lui dit des invectives, et il la précipite dans un puits. L'instruction du procès se poursuit devant les tribunaux ; et sur la déposition d'une foule de témoins, qui rappellent ses écarts emportés, il est condamné à une réclusion perpétuelle dans l'hospice des aliénés de Bicêtre. ” (*Traité*, III, §IX, p. 150)

Voilà donc le cas d'un sujet qui devient fou par “ nuances ”, cette folie étant – comme dans certains cas de manie – généralement accompagnée de calme apparent, et surgissant par à-coups. L'événement de folie est un acte criminel, qui fait de ce sujet alors considéré comme normal un fou – et précipite son internement. L'asile, ici, sanctionne le crime et en même temps est la conséquence d'une folie déclarée comme folie bien antérieurement, et manifeste seulement de façon intermittente : quelle que soit sa forme, l'aliénation doit finir

à l'hôpital, et la prison (dont le rappel des “ écarts emportés ” a protégé le sujet à son procès) n'est pas un terme approprié pour elle.

Comparons maintenant avec le cas suivant :

Cas F. “ Le médecin observateur peut remarquer quelquefois dans la société ce premier degré de démence dont on trouve des modèles complets dans les hospices. Un homme nourri dans les préjugés de l'ancienne noblesse, et à peine à sa cinquantième années, s'acheminait à grands pas, avant la révolution, vers cette désorganisation morale ; rien n'égalait sa mobilité et les aberrations de son effervescence puérile ; il s'agitait sans cesse dans l'intérieur de sa maison, babillait, criait, s'emportait pour les causes les plus légères, tourmentant ses domestiques par des ordres minutieux, ses proches par des inconséquences et des écarts brusques, dont il ne conservait un moment après aucun souvenir, aucune trace ; il parlait tour à tour avec la plus extrême versatilité, de la cour, de sa perruque, de ses chevaux, de ses jardins, sans attendre de réponse, et sans donner presque le temps de suivre ses idées incohérentes et disparates : une femme très spirituelle, que des convenances du rang avaient associée à sa destinée, tomba par cette union dans l'hypocondrie la plus profonde et la plus désespérée. ” (*Traité*, IV, §XVII, p. 161)

Ce cas montre que la folie en société ne conduit pas toujours à l'hospice ; mais il vise aussi à comparer la folie ordinaire en société de celle qui se vit entre les murs de l'asile. L'asile, apparaît comme l'étalon pour juger de la folie (“ remarquer quelquefois dans la société ce premier degré de démence dont on trouve des modèles complets dans les hospices. ”). En comparant avec le cas précédent (E) on voit que la différence est précisément le fait criminel propre à celui-là : la connexion de la société avec l'asile passe par la médiation du judiciaire. — Si on veut nommer “ médicalisation de la folie ” le processus dont on sait qu'il fera, cinquante ans plus tard, de la folie une maladie mentale objet du regard d'un expert (le psychiatre) résidant dans son lieu propre (l'asile), alors la lecture de ces deux cas démontre qu'elle ne va pas sans une complexe articulation du judiciaire et du médical, laquelle fait totalement défaut dans les cas de la *Médecine clinique*.

2.3.2. *L'hôpital et ses buts.*

Le régime pinellien du cas d'aliénation comporte donc une référence nécessaire de la folie à l'hôpital, même lorsque le fou n'est pas hospitalisé, et même quand il n'existe pas

parce qu'il n'est qu'un personnage de littérature comme le démontre éloquentement la présence du Ménélope de La Bruyère dans le cas F.. Pris dans cette référence commune à l'asile, tous les cas de folie obéissent à la même syntaxe : le cas F, comme les autres, comporte ainsi accumulation, vague référence causale permettant un semblant de compréhension pour toute personne au fait de l'histoire sociale (ici, le rapport ancienne noblesse – révolution, récurrent comme grille d'interprétation dans le *Traité*), et enfin, chute (non pas l'internement, mais une sorte de contagion: l'épouse devient hypocondre) - une chute qui seule justifie rhétoriquement l'arrêt du récit. A quoi sert donc cet asile ?

A lire les cas, à les réinsérer dans la conception pinellienne de la folie, on comprend qu'il vise avant tout à donner une prise sur cet être en quelque sorte profond du sujet, imperceptible la vue du profane, et sous-tendant sur la longue durée les manifestations parfois rares de folie. Le cadre de l'hôpital permet cette prise, par des moyens humains, bien sûr, mais aussi par l'isolement, surtout par la maîtrise du temps, enfin par un jeu entre artifice et réalité, dans lequel l'artifice est censé venir secourir, soutenir et reconstruire le sens de la réalité. En voici un indice : un insensé (cas N) donne " libre cours à ses actes d'extravagance ", il ne cesse de promettre mais recommence toujours ses " écarts insensés " : la quatrième " explosion " fit sentir au surveillant la nécessité de produire sur cet aliéné une impression durable et profonde. " (*Traité*, p. 101) Le médecin comme simple individu peut toujours obtenir des promesses qui s'avéreront illusoires, il faut donc pour qu'elles puissent être fiables et emporter la guérison, faire " sur l'aliéné une impression durable et profonde " : l'appareil de l'asile sert à cela.

Le but, c'est finalement d'obtenir *une parole* ; la parole de l'aliéné doit être une parole tenue. On sait que paroles sont fondamentales en société, que tenir sa parole est – sous le régime politique des théories du contrat, dont Pinel est le presque contemporain – la condition minimale de la socialité : le fou guérira donc en réintégrant donc cette faculté sociale minimale. C'est pourquoi elle est l'enjeu du processus, comme l'illustre cet autre cas, celui d'un militaire furieux : on le laisse " exhaler sa fougue impétueuse ", jusqu'à ce qu'il entrevoie lui-même qu'il " n'est pas le maître de suivre ses caprices ". Alors il s'adresse au surveillant " du ton le plus soumis " :

" Tu m'as promis, lui dit-il, de me rendre la liberté dans l'intérieur de l'hospice si j'étais tranquille ; eh bien, je te somme de tenir ta parole. " L'autre lui exprime, en

souriant, le plaisir qu'il éprouve de cet heureux retour sur lui-même. " (II, §VII, p. 59)

L'asile s'impose alors, à travers ces cas, comme le cadre ou l'appareil susceptible de produire, non pas tant l'aveu – comme les analyses de Foucault y ont longuement insisté³⁴ – mais la *parole fiable*. Rhétoriquement cette parole-but, produite par les moyens de l'asile, s'oppose alors à la parole-symptôme, la parole délirante, qui peut marquer l'entrée en folie, et qu'il faut toujours citer lorsqu'on relate le cas : c'est, dans notre cas A, l'allusion à Danton : " Comment, Danton un traître ! répète-t-il sans cesse ".

L'hôpital apparaît alors comme une machine à produire ce qui dans l'exemple N induit la capacité à promettre, à savoir " des impressions profondes ".³⁵ De même, "l'impression d'une crainte vive et profonde" doit frapper l'homme qui refusait de s'alimenter, dans la scène du potage (cas B). Peu importe la nature de cette impression, l'impression comme telle étant dans le cadre conceptuel de Pinel, aussi bien physique que morale, puisque la vie biologique est essentiellement sensibilité, pour Pinel comme pour Cabanis ou avant eux Bordeu³⁶ et Barthez³⁷. L'impression, dans la vie ordinaire, est légère, à moins d'un hasard improbable : d'où ces cas de guérison par un événement hasardeux, selon une sorte d'équilibre contradictoire des affects qui illustre un certain schéma de compréhension de la folie. Pinel a pu trouver ce schéma dans la tradition, et l'a justifié dans la physiologie de l'époque, pour laquelle l'économie animale est faite d'actions et réactions, comme le dit Ménuret dans l'article homonyme de *l'Encyclopédie*³⁸. Voici le cas :

Cas J. " Un homme de lettres, sujet à des excès de table, et guéri depuis peu d'une fièvre tierce, éprouve vers l'automne toutes les horreurs du penchant au suicide, et souvent il balance avec un calme effrayant le choix de divers moyens

³⁴ *Histoire de la folie*, p. 160.

³⁵ Le mot revient à plusieurs reprises : " l'impression produite sur l'esprit de l'aliéné a été des plus profondes " (*Traité*, p. 236, cf. cas H plus bas).

³⁶ Voir sur ce point Elisabeth Williams, *A Cultural History of Medical Vitalism in Enlightenment Montpellier*, London, Ashgate, 2003 ; Wolfe et Terada, *art. cit.*, Huneman, "Passions".

³⁷ Paul-Joseph Barthez, *Nouveaux éléments de la science de l'homme*, 2nd ed., 2 vols, Paris, Goujon & Brunot, 1806 ; Théophile Bordeu, *Recherches anatomiques sur la position des glandes et leur action*, Paris, Quillau père, 1751.

³⁸ Sur économie animale, action et réaction, voir : Wolfe et Terada, *art. cit.*, Huneman, " Passions » ; Starobinski, *Action et réaction. Vie et aventures d'un couple*, Paris, Seuil, 1999, p. 107sq.

propres à se donner la mort. Un voyage qu'il fait à Londres semble développer avec un nouveau degré d'énergie sa mélancolie profonde et la résolution inébranlable d'abrégier le terme de sa vie.³⁹ Il choisit une heure très avancée de la nuit, et se rend sur un des ponts de cette capitale pour se précipiter dans la Tamise ; mais au moment de son arrivée des voleurs l'attaquent pour lui enlever toutes ses ressources, qui étaient très modiques ou presque nulles ; il s'indigne ; il fait des efforts extrêmes pour s'arracher de leurs mains, non sans éprouver la frayeur la plus vive et le plus grand trouble. Le combat cesse, il se produit à l'instant une sorte de révolution dans l'esprit du mélancolique ; il oublie le but primitif de sa course, revient chez lui dans le même état de détresse qu'auparavant, mais entièrement exempt de ses projets sinistres de suicide. Sa guérison a été si complète que, résidant à Paris depuis dix ans, et souvent réduit à des moyens d'existence précaires, il n'a plus éprouvé le moindre dégoût de la vie. C'est une vésanie mélancolique qui a cédé à l'impression de terreur produite par une attaque imprévue. ” (V, §VI, p. 242)

Bien sûr on ne saurait généralement compter sur ce genre d'événements. L'asile vise donc à les reproduire de manière contrôlée, à retenir le mécanisme susceptible d'impressionner profondément – ce que ne sauraient faire ni la simple “ direction d'esprit ” dans la tradition religieuse ou philosophique, ni les drogues – sans le soumettre aux vicissitudes du hasard. Tel est le sens du cas suivant (cas H). Un ouvrier, lors de la Révolution, se croit menacé de mort suite à des propos critiques qu'il aurait tenus. “ L'idée d'être condamné à périr par la guillotine l'absorbe tout entier nuit et jour ” (*Traité*, 234) Pinel invente alors un “ expédient ” pour guérir cet aliéné transféré à Bicêtre : il charge trois jeunes médecins de constituer une fausse Commission du corps législatif, qui va examiner son cas, délibérer et finalement prononcer un acquittement. “ On se retire en silence et tout indique que l'impression produite sur l'esprit de l'aliéné a été des plus profondes. ” (*ibid*).

Une telle “ impression profonde ” provoquée par la simulation d'un événement fort ne saurait évidemment subsister si elle n'est point soutenue : “ sans action ” (comprendre : sans travail), notre ouvrier rechute ; et, apprenant par hasard que l'événement fut une

³⁹ Pour l'époque, Londres est la capitale de ce pays connu pour *The english malady* depuis George Cheyne (*Natural method of cureing the diseases of the body, and the disorders of the mind depending of the body*, London, Strathan, 1742) c'est-à-dire un mal moral ; le suicide des Anglais, suicide curieusement endémique et ayant pour particulière de n'avoir aucune cause compréhensible, est un lieu commun pour les lecteurs de Pinel ; cette phrase est donc narrativement signifiante.

manigance, sa rechute est irréversible. Par la négative, ce cas nous enseigne donc deux choses : d'abord, “ “ l'unique garant du maintien de la santé, des bonnes mœurs et de l'ordre est la loi d'un travail mécanique rigoureusement exécuté ” (*Traité*, p. 224), et ensuite, l'asile doit être isolé, car l'aliéné ne doit pas être en communication avec l'extérieur, sans quoi ces communications précisément livrées au hasard des caprices humains risquent de saboter l'effet de l'asile. D'où l'interdiction de contacts prolongés avec la famille, dont un autre cas vient illustrer l'effet néfaste :

Rhétorique du cas clinique chez Pinel

Le cas n'est donc pas seulement informatif, son récit n'a pas une simple vertu cognitive ; il est au service d'une cause, il veut démontrer l'importance du traitement moral et la nécessité du cadre asilaire. C'est pourquoi Pinel parmi les premiers cas commençait par celui de l'ami, à qui l'asile fit défaut. Le *Traité* de 1802 est un livre militant, avec trois thèses-cibles : il expose la nature de l'asile comme lieu de guérison et en revendique la nécessité ; corrélativement il pose que la folie est essentiellement curable ; et que cette cure relève d'une procédure spécifique, le “ traitement moral ”, apanage d'un type particulier de médecin, l'aliéniste. Il s'ensuit alors une fonction du cas, qui n'est pas seulement de donner à comprendre la nature de la folie, le principe de sa division en espèces, mais dans le même temps de justifier les revendications pinelliennes. Ainsi, le cas N, déjà cité, illustre la curabilité de la folie :

“ Un père de famille, très recommandable, perd sa fortune et presque toutes ses ressources par des événements de la révolution, et une tristesse profonde le conduit bientôt à l'état de maniaque. Traitement routinier et ordinaire de la manie par les bains, les douches, les saignées répétées, et les moyens de répression les plus inhumains ; les symptômes, loin de céder, empirent, et on le transfère à Bicêtre, comme incurable. ” (*Traité*, p. 100)

Ce parcours est à l'époque ordinaire : Bicêtre recueille les “ irrécupérables ”, et Pinel, son directeur, est donc en position stratégique pour faire valoir la curabilité de la

folie contre les arguments courants. Ainsi, “ sans s’arrêter aux avis qu’on lui donne en désignant cet aliéné comme très dangereux ”, le surveillant, à savoir Pussin, “ le livre un peu à lui-même pour étudier son caractère ”, et ce patient finira par guérir.

Raconter le cas est donc une opération rhétorique : non seulement informer et convaincre, mais aussi obtenir des effets par ce discours. Cela est crucial pour la narration des guérisons, qui sont déjà, dans l’asile, souvent des mises en scène, mais dont le récit obéit lui-même à certaines règles. Les cas, généralement, ont une chute : la guérison, lorsqu’elle intervient par une procédure typiquement asilaire, est une pointe, au sens presque rhétorique que le mot prend chez La Rochefoucault. Voici donc pour finir un cas exemplaire :

Cas M. “ Un des plus fameux ouvriers de Paris s’infatue de la chimère du mouvement perpétuel et, pour y parvenir, il se livre au travail avec une ardeur infatigable ; de là la perte du sommeil, l’exaltation progressive de l’imagination, et bientôt un vrai délire par le concours des terreurs renaissantes qu’excitaient les orages de la Révolution. Le renversement de sa raison est marqué par une singularité particulière. Il croit que sa tête a tombé sur l’échafaud, qu’on l’a mise pêle-mêle avec celles de plusieurs autres victimes, et que les juges, par un repentir tardif de leur arrêt cruel, avaient ordonné de reprendre ces têtes, et de les rejoindre à leurs corps respectifs ; mais que, par une sorte de méprise, on avait rétabli sur ses épaules celle d’un de ses compagnons d’infortune. L’idée prédominante de ce changement de tête l’occupe nuit et jour, et détermine les parents à lui faire subir le traitement des maniaques à l’Hôtel-Dieu ; il est ensuite transféré à l’hospice des aliénés de Bicêtre. Rien n’égale alors son extravagance et les éclats bruyants de son humeur joviale. Il chante, il crie, il danse, et comme sa manie ne le porte à aucun acte de violence, on le laisse errer librement dans l’hospice, pour exhaler cette effervescence tumultueuse. (...) Vers l’approche de l’hiver ses emportements s’apaisent (...) L’idée du mouvement perpétuel se renouvelle au milieu de ses divagations insensées ; il crayonne sans cesse sur les murs et sur les portes les dessins du mécanisme propre à l’opérer. Comment l’arracher à cette chimère, sinon par l’inutilité de ses efforts multipliés et une sorte de satiété ? On engage les parents à envoyer quelques outils d’horlogerie (...) Le surveillant de l’hospice (...) lui permet de dresser une sorte d’atelier dans son antichambre (...) encore quinze jours d’une application soutenue, il rassemble alors toutes ses pièces, croit y voir un accord parfait, d’autant mieux qu’il en résulte un mouvement qu’il juge propre à se reproduire (...) Les rouages s’arrêtent, et le prétendu mouvement perpétuel ne dure que quelques minutes.

La confusion succède à l'ivresse de la joie (...) il déclare qu'il pourrait facilement lever l'obstacle mais que, fatigué de ces essais, il ne voulait plus s'occuper d'horlogerie. Il restait encore une idée délirante à combattre et à ruiner, c'était celle de son prétendu changement de tête, qui d'ailleurs se renouvelait au milieu de ses travaux. Une plaisanterie fine et sans réplique parut propre à l'en corriger. On, prévint un autre convalescent, très plaisant et d'une humeur gaie, du rôle qu'il aurait à jouer, et on lui ménage un entretien suivi avec l'artiste ; cet autre tourne adroitement le propos sur le fameux miracle de Saint-Denis, qui, chemin faisant, portait sa tête entre ses mains, et ne cessait de lui faire des baisers. L'horloger soutient fortement la possibilité du fait, et cherche à le confirmer par son exemple propre. Son interlocuteur pousse alors un éclat de rire, et lui réplique avec un ton moqueur : " Insensé que tu es, comment Saint-Denis aurait-il pu baiser sa tête ? Etait-ce avec son talon ? " Cette réplique inattendue et sans réponse frappe vivement l'aliéné ; il se retire confus, et il n'a plus parlé désormais de son changement de tête. " (*Traité*, II, §XII, p. 67 sq)

La guérison, spontanée, événementielle, fait ici contrepoint à l'origine mythique, événementielle elle aussi ; elle est une parole, qui met le sujet face aussi bien à la réalité (il a une tête) qu'à sa folie. Le récit du cas est tendu entre ces deux événements mythiques, chuter dans la folie et s'en relever, entre l'origine de la folie comme événement et la guérison instantanée comme pointe. Cette constitution générale, exemplifiée à la perfection par ce cas précis, donne son originalité au cas psychiatrique, puisque le cas clinique simple en est dépourvu, étant quant à lui le récit du processus plus ou moins sinusoïdal de la guérison ou du décès. Le récit donne alors à voir dans sa construction même la construction du cas comme cas psychiatrique : d'un côté une reconstruction mythifiée de l'origine : le sujet travaille sur le mouvement perpétuel et devient en quelque sorte prisonnier du mouvement, le déni de l'irréalité de ce mouvement étant l'expression même de son déni de la réalité ; de l'autre la construction du dispositif assurant la possibilité et l'efficacité de la parole qui résout. Dans l'exposé même de ce cas est donc mobilisé tout un dispositif qui met en jeu médecin, aides, moyens matériels, et même complicité d'autres malades (" un convalescent "), donc une certaine communauté psychiatrique.

Ce cas est éminemment signifiant si l'on prend garde aux connotations : voilà un homme qui, comme on dit, a perdu la tête, et qui, au lieu de le savoir et de le dire, met en acte cette vérité de lui-même qu'il ne sait pas. Le fait de dire " au sens propre " qu'il a

perdu la tête indique qu'il a perdu la tête " au sens figuré". Il s'agit finalement de le rétablir dans l'usage du langage, d'inverser cette dichotomie du propre et du figuré que la folie vint dérégler. Pour ce faire, on le met dans une situation où le langage, seul, produira cet effet. Ce cas-ci est complet : cause , soit une origine à la fois convenue (le travail de l'esprit rend fou, c'est bien connu, les cas C et M avaient recours à ce *topos*) et inexplicable ; internement ; prise en charge du délire dans une sorte de dispositif destiné à le contrer, à le mener à une sorte d'épreuve de réalité : le mouvement perpétuel est impossible ; enfin pointe finale : la parole qui confronte le sujet à son délire. Cette perfection rhétorique donne à voir que la folie est curable, et surtout, que la parole produit des effets propres à la guérir, à condition de ne pas vouloir raisonner le fou mais d'émerger comme la pointe d'un dispositif complexe qui désarme le délire. Le langage du fou est dérégulé – croire avoir perdu sa tête, c'est avoir perdu la tête - si bien qu'il faut le re-régler, donc parler au fou ; ce qui suppose que le fou est accessible à la parole.⁴⁰ Et en même temps, il n'est pas accessible à n'importe quelle parole, et surtout pas au discours habituel, symétrique, spontané et surtout sans but déterminé autre que de dire vrai : cette parole-ci doit être instituée d'une manière précise, émerger d'une configuration (insue du fou) que seul un asile peut créer de manière systématique - en particulier, en protégeant une telle parole des tiers susceptibles d'en désamorcer les artifices. La reconstruction en quelque sorte spectaculaire de notre cas M expose de la manière la plus achevée le cas clinique dans le *Traité médico-philosophique sur l'aliénation mentale* : la narration d'une procédure et d'un dispositif spécifique de langage, apte à produire une certaine manifestation de la parole comme signe authentique de la guérison – narration qui elle-même, en tant que discours, doit produire des effets sur notre image de la folie et sur notre conception de son traitement.

Conclusion

Il n'y a pas de théorie à proprement parler du traitement moral, non seulement parce qu'on ne dispose pas d'une théorie causale de la folie et donc de sa cure, mais parce que cette absence est la condition même, en quelque sorte, de la pratique psychiatrique. Le

⁴⁰ Comme y ont insisté les belles analyses de Gladys Swain, *op. cit.*, esp. p. 100 sq.

traitement moral s'institue dans l'espace ouvert par l'ignorance de la cause de la folie, et la décision de négliger cette cause et l'investigation de sa nature, de ne pas se commettre dans l'assignation de la folie à une cause organique ou à une justification morale – d'où la section du *Traité* qui réfute l'attribution d'un siège à la manie – : on a vu combien lapidairement la cause était traitée, ou combien mythiquement elle était reconstruite. Mais précisément, la médecine aliéniste ne peut donc fournir dans ce traité *médico-philosophique* une théorie scientifique de la manie, ni même une justification purement théorique de sa classification des vésanies et de ses procédures thérapeutiques. Sa manière de procéder consiste donc à raconter les cas, mais à les raconter dans une grammaire proche du cas clinique médical type (d'où un effet de “ médicalisation ” à la lecture du traité). Surtout, il s'agit de les relater en mettant en évidence l'utilité du traitement moral, en tant qu'il est le seul qui corresponde à ce type de cas médicaux. Le cas est donc à la fois illustratif et essentiel parce qu'il représente, dans son écriture même, l'exigence d'autonomie de la médecine aliéniste. Pareille exigence se fait entendre parce que le cas, en lui-même, comporte un rapport nécessaire, sous divers modes, au régime hospitalier. Ce rapport a lieu en particulier, dans l'identification de la folie, dans la dialectique du caché et du manifeste, et enfin dans le processus par lequel le rapport du visible et de l'invisible dans la folie ne peut se constater que dans l'intérieur du rapport du médecin au malade, le médecin à son tour n'étant médecin effectif qu'à disposer autour de lui d'un “ hospice bien ordonné ” (plutôt que d'un savoir médical, d'une batterie de théories explicatives).

Le cas psychiatrique selon Pinel obéit donc à plusieurs exigences corrélatives : frapper l'œil ; rendre narrativement compréhensible un processus d'entrée en folie et de guérison alors même qu'il n'y a pas de théorie explicative de ce processus ; rendre sensible la carence d'asile et de traitement moral lorsqu'il n'y a pas guérison. Ceci s'opère par le biais d'une reconstitution plus ou moins mythique de l'origine causale et de la guérison, de manière à faire du cas non pas, comme dans la médecine clinique, un récit singulier correspondant à la complexité réelle d'une espèce pathologique faite d'espèces simples, et propice pour le médecin à reconstituer la nature de la maladie puis à l'enseigner au novice – mais bien une narration reconstruite selon d'autres règles, celles de la rhétorique de *l'origine* et de la *pointe*, selon une temporalité reflétant la temporalité longue de l'hôpital.

Le cas de la médecine clinique est prélevé sur la masse de malades de l'hôpital ; en l'individualisant, sa description permettait d'individualiser la maladie en analysant les espèces dont elle est constituée. Il a l'individualité du cas par rapport à la règle, individualité que parfois marque le nom de l'individu ; du côté du sujet, le langage n'est point troublé : celui-ci se nomme et nomme son mal. Au contraire le fou est toujours anonyme, mais son langage l'individualise comme sujet malade. Sous des espèces plus ou moins délirantes organisées en un continuum qui va de la manie sans délire à la démence et que le *Traité* nous aura exposé, ce langage fait déjà partie du cas psychiatrique, de sa manière d'être fou, il suffit à individualiser le cas. Et parce que le fou parle, mais que sa parole est dérégulée, la folie du fou se traite presque toujours avec une parole, mais une parole sous conditions spéciales, artificielle, parole produite par un dispositif qui bientôt deviendra le dispositif asilaire, seul à conjoindre " l'hospice bien ordonné " et " l'emprise d'un homme ", comme le *Traité* l'exige au départ.

Mise en scène de cette parole, tel est l'hôpital rêvé et revendiqué par Pinel ; pour ce faire, les cas sont comme la mise en mots, rhétorique, de cette mise en scène. Dans cette mise en mots d'une mise en scène, l'aliéniste, avec ses " qualités morales et physiques ", et les instruments de son " emprise ", est toujours là – au moins pour nous, et même lorsque le sujet aliéné se croit seul. Différence ultime avec le cas selon la *Médecine clinique* : ce dispositif constitué du médecin et de l'appareil asilaire autour de lui est toujours présent dans le cas lui-même, il représente un tiers qui empêche que le cas clinique soit la pure mise en relation d'une espèce pathologique et d'un individu humain : ce tiers qui en retour est susceptible à la fois de faire surgir la pathologie là où elle est invisible au profane (dialectique du caché et du révélé), donc précisément, d'" impressionner profondément " cette espèce de fonds, que Pinel se garde bien d'identifier comme psychique ou comme organique, et dont l'aliénation est censée provenir, avec toutes ses espèces identifiables comme telles par ce médecin dans l'enceinte hospitalière.⁴¹

⁴¹ L'auteur remercie chaleureusement Patrick Singy pour sa relecture impeccable et ses judicieuses suggestions.