Securing the Future of the Critical Project of Aufklärung

Rosalba Mallardo

Thesis submitted for the degree of PhD in Philosophy

School of Philosophy and Art History

University of Essex

January 2016

Contents

Abstract

Introduction ..... v-xiii

I. Habermas and the possibility of Discourse Ethics..... 1

Communicative Reason and Discourse Ethics..... 7
Sublation..... 14
Coexistence..... 24
Cooperation..... 34
Gattungsethik—an ethics of the species..... 54

II. Conjugating the Historical with the Fundamental..... 73

The Trans-historical and the Historical..... 79
Modernity and the Emergence of Radical Responsibility..... 83
The Articulation of Modern Critique..... 97
Radical Critique..... 106
Radical Self-Critique and the Limits of Interpretative Knowledge..... 117

III. Foucault and the Re-conceptualization of the Project of Modernity..... 141

The Trap of Modernity and the Constant Possibility of its *Ausgang*..... 151
The Archaeology of Knowledge and the Possibility to Think Otherwise..... 164
The Power to Disconcert..... 171
The Care of the Self..... 177
Answer to the main criticisms..... 185

Bibliography..... 226

Abstract

Modernity is stuck between the potential of its emancipatory spirit and the difficulty of implementing it.

This impasse has generated the need to look for a feasible form of critique that is capable of grounding reason in order to revitalize the modern project of emancipation. In this aim, the dominant strategies, intend to provide reason with a way to mediate its own comprehension: radical critique.

In this regard, I evaluate Habermas' and Taylor's work for they are two paradigmatic and opposing examples of that strategy. I come back to them because there are still generations of thinkers that use, or can be recognized as using, the philosophical instruments shaped by them. However those instruments are structurally faulty as I demonstrate, from new perspectives, in Chapters I and II, respectively.

Their failures, I argue, either open up the way to an idle repetition of the same mistakes or, (should) push us to look for an alternative approach to the modern project of emaciation as there is still the need of it.

Against the impracticable option of using ill-suited tools, without having explored alternative approaches, I investigate what I take to be a re-conceptualization of the modern project of emancipation.

Through Foucault's work, with the Chapter III, I explore and then defend a project that I reckon as capable to both explain why radical critique fails, and open up a feasible new path for emancipation.

Introduction

With the decline of the old metaphysical order we, the moderns, could have expected to finally benefit from an effective, radically critical reasoning—a reason that no longer accepts anything but what passes the scrutiny of reason itself.

The liberation from the burden of all those elements that did not pass that scrutiny should have caused the dissolution of myths and superstition, the withdrawal of religion into the private sphere, and the strengthening of the natural sciences. Such a process, understood as progressive and irreversible, should have been able to produce an emancipatory power which, in turn, should have led us toward a triumphant and universal freedom.

Yet things did not go exactly like that. Modernity, since its very beginning, revealed itself as the locus of much more ambivalent processes.

The old metaphysical order, though shattered, did not simply disappear. Nor did it survive as an idle and ineffectual wreck. On the contrary, its fragments endure within modernity without a readymade legitimation. Simultaneously, reason—turning on itself—has become also the foundation, and not merely the instrument and the object of self-comprehension. This reflective turn, which has inaugurated the path of Enlightenment, has, however, also revealed the dangers of a radically reflective reason. As a foundation, modern reason, while allowing for a maximization of its negative critical function, has been unable to comprehend itself and to forge, on that basis, a reliable way to justify the aims it legislates. The absence of this reliable foothold has caused a dilemma concerning the

manner in which reason ought to determine its own aims. In particular, this shortcoming risks undermining reason's ethical function of providing a positive practical orientation.

At the same time, instrumental reason has been flourishing. With the intent to increase both knowledge and freedom, virtually every level of reality has been subject to possible or actual processes of rationalization. In the face of enormous advancement—from the re-organization of the public life, to the specialization of knowledge—those processes have as direct and/or side effects also phenomena of uncritical automation, abstraction, and ultimately alienation.

In other words, the new modern paradigm, despite the liberation of reason, has also triggered dysfunctional processes that constantly betray and hinder the modern project of human emancipation.

In the light of this problematic condition, the comprehension of what has gone wrong with modernity, and how to correct and revitalize its emancipatory spirit has become the essential task for modern philosophy; and consequently critique, as potentially able to provide an answer to this question that modernity raises, has acquired a renewed and peculiar centrality.

Critique, at the epistemological, ontological, and socio-political level, is supposed to assess, and if necessary, correct the aims that reason itself endorses; and at all these levels, critique must operate on an unstable foothold—modern reason itself. This means that it is not enough to simply understand how appropriately the aims that reason gives to itself are implemented—simple critique—but also, and above all, what is central to modernity is the necessity of questioning those very aims reason gives to itself and the way it justifies them—radical critique.

Yet, the presence of a necessity, even when as enduring as the question of modernity is, does not itself suggest the existence of an effective answer. In fact, the account of modern reason, that, through its own functioning, risks of betraying its emancipatory spirit, starting from the end of the nineteen century, has been a recurrent theme described, with specific variations, by many—Weber, Adorno, and Horkheimer ranking among the most influential.

While all their views about the fate of modernity have been, in different ways, pessimistic about the possibility of answering the question of modernity and therefore of realizing the emancipatory project of our time, the second half of the twentieth century has also seen renewed attempts to provide a more stable foothold for modern reason, thus showing the modern project in a more positive light.

In these renewed attempts, it is possible to distinguish two competing approaches that we could define, very broadly, as Kantian and Hegelian; not because those approaches are pedantically faithful to those authors, rather, because such a definition marks the characteristic source they respectively drawn on in order to shape a viable form of radical critique.

While the Kantian approach sees the conditions of validity of radical critique as ingrained in the formal and universal structures of reason itself, the Hegelian approach recognizes in history, *qua* learning processes, the indispensable framework for the functioning of a valid form of radical critique.

In what follows, I look at two contemporary and paradigmatic expressions of these two approaches.

Jürgen Habermas and Charles Taylor not only provide us with insightful analyses of both the historical and theoretical condition of modernity, but they also explicitly present

their own works as continuing the project of modernity. For both recognize, even if from different perspectives, that it is because of the plight of modernity itself that radical critique has became the central task for philosophy.

Though, as I show in the first and the second chapter of the present work, neither of the two, for structural reasons, is ultimately able to realize the demanding goal they set for themselves and for modern philosophy as a whole.

In spite of their failures, the analytical instruments they have shaped, the theoretical perspectives from which they look at the modern malaises, and in general their strategies have been adopted by, or can be seen in, a large and heterogeneous cluster of contemporary thinkers.[1] And because these thinkers are adopting, methodologically speaking, approaches relevantly similar to those of Habermas and Taylor, we now encounter a new problematic stasis—a curious phenomenon.

Unconcerned by the dangers related to the incapacity of their strategies and instruments to help us resolve the impasse of modernity, those thinkers have constructed within that very impasse their own houses.

We can make sense of this curious phenomenon in two different ways: we can understand it either as a mistake or as a crypto-renounce. Accordingly, I see the cluster of thinkers I am referring to as divisible between two broad camps.

The first (i) must still be ignorant of the fact that their approaches are, structurally speaking, unable to ground the project of modernity—that is to say a feasible version of radical critique. The second group (ii) while presenting themselves as endorsing critical

[1] Liberal theorists as Reiman (1990), Scanlon (1998), in some extent Rawls himself, Benhabib (1992), and, as expression of the second approach, Raz (1999), Sandel (2007), and Walzer (1980), only to mention some among the most influential.

projects, no longer aims at shaping a stable foothold for modern reason and radical critique.

The first group can be described as continuing a project without the means that are fit for purpose. It is for this reason that subjecting the approaches—approaches developed in a paradigmatic way by Habermas and Taylor—to further criticism, but from novel perspectives, can be justified; for it is by showing this mismatch between means and aim that I aim to convince them that the way in which they continue to engage with the critical project of modernity is problematic.

However, I reckon that the position of the second group is even more problematic, because the problem is not so much a mismatch of means and goals, rather it is that the failure of the modern project seems to be taken for granted, while not wholly admitted. At the same time, though, members of this group still use the philosophical approaches created as means to get the project of radical critique off the ground.

This inconsistency between the actual practice and implicitly giving up the aims of radical critique turns philosophy into an intellectualistic pose which is not only not useful, but actually dangerous. The implicit manner in which the end of the full scope of the project of modernity is stated actually prevents the very possibility of engaging that statement in a direct and fruitful manner, since the focus of the debate is always elsewhere. And it is only against the backdrop of such a debate that we could decide whether continuing to work with such approaches is philosophically justifiable or not.

In the meanwhile, the condition of modernity, and all the malaises identified by the first generation of critical theorists, are not only all still present, but they have become, if anything, even more intense as, for instance, the hyper-bureaucratization of national and supranational institutions to the detriment of the political attests. This means that the

necessity of revitalizing the project of modernity, or the exploration of whether this project can be re-interpreted, has not faded away.

It is for this reason that, with the third chapter, I look into what I take to be a feasible way of revitalizing the emancipatory project of modernity by drawing on the work of Michel Foucault.

By analyzing modernity anew, he goes back to the original intention to build a foothold capable of securing the functioning of reason, and, in the very spirit of modernity, calls (the dominant interpretation of) that task into question. I interpret his work as a re-conceptualization of the project of modernity and not a way of giving it up. I will call his approach a form of hyper-critique as it questions the traditional way of understanding the question of modernity and shows that it gives rise to a feasible continuation of this project that does not require us to embrace any defeatism concerning emancipation.

Chapter I is devoted to Habermas' endeavor to revitalize the project of modernity through a functioning form of radical critique. First I reconstruct the fundamental elements of his formal pragmatics, stressing the fact that such form of radical critique is meant to be both receptive to the conditions of modernity, and autonomous enough to support universal validity claims. I then test the feasibility of his project by considering two case studies.

The test consists in (i) investigating whether Habermas' radical critique is able to assess religious content, that is, a pre-modern fragments that one would imagine as not particularly problematic for modern reason to assess; and (ii) evaluating whether Habermas can deal with an attack on what Habermas himself understands as condition of existence of his entire critical project — i.e. an attack on species ethics.

While the first case study unfolds chronologically, following Habermas in what I interpret as his evolving attempts (sublation, coexistence, and cooperation) to use, without succumbing to, the potential motivational sources he individuates in the Judeo-Christian tradition, the second case study focuses on the sole text Habermas devoted to the concept of species ethics—*The Future of Human Nature.*

What emerges from these two studies is that Habermas' endeavor fails regarding its own aims. It fails, in both cases, because his formal pragmatics, instead of being merely receptive to, actually depends upon features of the facticity it was supposed to validate, and such a dependence invalidates any radically critical project.

The elements against which Habermas' project crashes are integral parts of Taylor's own attempt to build a reliable foothold for modern reason. Taylor, in fact, approaches the question of modernity by beginning from what already matters to us moderns.

He reckons as possible the shaping a form of radical critique that draws on interpretative knowledge, for according to him, a reflective judgment is articulated, and cannot avoid being articulated by starting from a minimal self-knowledge of the individual involved.

Building on this, Taylor explains that this knowledge which constitutes the minimal condition of possibility of a reflective judgment is, in turn, made possible by the interplay of two elements: (i) trans-historical and immutable formal conditions such as the sense of the self, as well as coherence and unity of experience, and (ii) the historical moral sources (as the conceptions of the good) in which the identity of each is embedded.

Taylor's strategy, which I reconstruct in Chapter II, consists, firstly, in turning self-interpretation into a reasonable foundation for the form of radical critique that emerges from the interaction of (i) and (ii). Taylor then aims to show, through paradigmatic

examples, how the implementation of such a form of self-interpretation can be considered both valid and radically critical, for it allows a form of reasoning which is based on transitions and which can account for the progressive character of those transitions.

I show that Taylor's attempts fall short of his ambition to underpin a radical form of critique since it is either drawing on a naïve understanding of history, or, if rescued from that danger, then, it is unable of dealing with the integration of new forms of significance within what is already authoritative.

Against the backdrop of the failure of Habermas' and Taylor's paradigmatic attempts of revitalizing the project of modernity by grounding radical forms of critique, I move on, in Chapter III, to Foucault's re-conceptualization of the project of modernity.

I present Foucault's work as both a critique of the epistemological limits of radical critique, and the attempt to re-think critique itself as a practice endowed with analytical tools—therefore not extraneous to the highest standards of modern philosophy—yet not fully justifiable on normative and epistemological grounds.

I start the analysis of Foucault's work exactly from the definition of such practice. He, appropriating the concept of *Aufklärung*, describes it as an ethical disposition, and then proceeds to qualify this disposition as the circumstantial will of not being subjected to specific games of knowledge and power—hyper-critique.

Subsequently, I describe archeology, genealogy, and care of the self as the tools shaped in order to practice and sustain such a form of hyper-critique.

Finally, I provide a defense of the feasibility of Foucault's critical analytics against the main criticisms that have been raised, and in particular against the one of reproducing, within his own work, the fallacious logic he identified at the ground of radical critique. In

doing so, I present an argument to the effect that his continuation (cum transformation) of the modern critical project is superior to the alternatives presented by his competitors.

Habermas and the possibility of Discourse Ethics

Habermas attempts to formulate both (i) an analysis that is able to depict the processes that are preventing modernity from developing accordingly to its full emancipatory potentiality, and (ii) a pragmatic method that could finally provide modern reason with the anchoring that it needs in order to implement such potentiality. In doing so he places himself in a tradition that, with respect to point (ii), grew increasingly pessimistic.

The first generation of the Frankfurt School, and the authors who orbited around it, all agreed about modernity being afflicted with a disease whose cure was as needed as difficult, or even out of reach. Horkheimer's *Eclipse of Reason,* in particular, epitomizes the analysis of a reason that, once it has lost its connection with its objective intelligible foundation, cannot but overdevelop its functional aspects, becoming increasingly abstract, formalistic, merely calculating, and ultimately irrational.

Horkheimer identifies a sort of autoimmune disorder: the very system of rational defence, that is the negative and destructive power of critique, ends up destroying reason as ethical agency. Once any objective, and therefore legitimized foothold, is shattered, then means and aims can be valued only on the ground of their efficiency, and every kind of speculation which tries to go beyond mere functional reasoning, given the very absence of a foothold, becomes either meaningless or banally whimsical.[1] The problem according to Horkheimer is not the fact that the intellect in its critical modality destroys what, in his terminology, is called false objectivity, that is to say all those elements which are not or no longer considered valid. That is the exact role that the critical element of reason is

[1] Max Horkheimer, *Eclipse of Reason*, New York, Oxford University Press, 1947, pp. 18 and 34.

supposed to play in relation to the objective one. Critique ought to call into question what has been considered valid, just, and ultimately legitimate in its existence. According to Horkheimer, the real catastrophe is that the critical activity of reason—the modern subjective reason in his words—despairs of "developing an objectivity of its own".[2] From the destruction of the old order no new objective landscape emerges. Modern reason, in its most formalized stage, negates not only its own objective articulations, which are constituted by all those contents that were once part of the pre-modern order, but it has also lost the ability to determine its new true ends. The consequences of this are terrible according to Horkheimer. Without objectivity, reason does not have any foothold to mediate its own comprehension, to allow for a positive use of critical power, and to ponder prospective modifications.

Despite Horkheimer's cry for a renewed role for critique, he, in the *Eclipse of Reason,* has no instruments to positively address this issue. What emerges from the text, more than a real answer, is rather the risk of getting stuck in a state of perennial ecstatic negation that makes it impossible to correct the reciprocal (dis)equilibrium of reason's elements, precluding, in this manner, the forging of a relation with self, nature, and others neither regimented under the banner of any ideology, nor characterized by functional domination.

Horkheimer's hint for a strategy is to entrust critical theory to the spontaneous reactions of reason—reactions to the painful symptoms derived from the unbalanced condition and functioning of reason itself. This idea, though, is unable to go further than the mere negation of the symptom. It, in fact, leaves completely untouched the question that Horkheimer himself considered as central, that is to say, how reason can possibly

[2] Ibid., p. 7.

revitalise the power to produce its own valid objectivity. And not only the question is left unsolved, but the entire work ends with Horkheimer himself reinforcing the pessimism of the title concerning the actual chances of modern critique to be revitalized.

It is within this *impasse* that Habermas starts working, and yet his approach will not be so dark.

Habermas shares the diagnosis, which is almost entirely absorbed within his own reconstruction of the malaises of modernity, but does not share the pessimism which characterizes not only Horkheimer's work, but the whole first generation of the Frankfurt School. Contrary to that aspect of said tradition, Habermas thinks that a new form of objectivity is not only urgently needed but possible.

Habermas believes that, in particular, Horkheimer's pessimistic stance derives from two main elements: he did not understand the irony of the entire modern process, and he also ended up identifying completely and erroneously rationalization with reification. The missed irony and the following erroneous overlapping consist in not seeing that the potentiality for a non-oppressive unconditionality, that is to say, the new and much needed form of objectivity, could find space only within the crisis that the modern form of instrumental reason has brought about in the so called western societies. This potentiality that before the Enlightenment was caged by the fixity of worldviews, has been, due to the autonomous formalizing forces of modern rationality, first liberated and then partially distorted. Despite these distortions, Habermas thinks we can draw on these liberated potentialities and reckons it is possible to solve the dilemma that reason used against itself has caused. Free and democratic opinion, unfastened from the old, and no longer valid

objectivity, had to emerge so that an unprecedented form of reason could be finally put to work.[3]

Even if Habermas had clear in his mind since the second half of the 60s the kind of theoretical path he intended to walk in order to answer the question of modernity, the detailed comprehension of the condition of modernity, and the elaboration of the theoretical tools to actually bring about and secure the new fragments of uncoditionality took a long time.[4] The answer Habermas was looking for had to meet precise criteria.

After the destruction of the metaphysical order, any critical theory that aims to, and has to, use the residues of that destruction – since this is the only available material – must find a way to justify itself while already being at work. What is needed then must be pragmatic in method: something able to work without relying on *a priori* categories. But, at the same time, it must be able to scrutinize any possible content; that is to say, Habermas' answer must have also a universal character—and this in two distinctive ways: 1) the access to the critical resources, and the methodology employed must be sharable by every rational subject; and 2) the output of this scrutinizing process must have a claim to be universally valid. In other words, the fact that Habermas' critical project must be necessarily receptive to elements already historically available, does not mean that it must renounce to be intrinsically rational and normative.

What is more, the reflection on the status of modern rationality brings in the picture another fundamental aspect: the so called linguistic turn. With the destruction of the monological order that shaped the pre-modern time, the subject and its consciousness

[3] Habermas, *Theory of Communicative Action*, trans. T. McCarthy, Boston, MA, Beacon Press, 1984/1987, Vol. II, pp. 332-334. From now on TCA.

[4] Already in the 1965, with his inaugural lecture at Frankfurt University, Habermas states: "What raises us out of nature is the only thing whose nature we can know: *language*. Through its structure, autonomy and responsibility are posited for us. Our first sentence expresses unequivocally the intention of universal and unconstrained consensus." See Habermas, *Knowledge and Human Interest,* J. Shapiro trans., Beacon Press, Boston, 1968, p. 314.

underwent a significant transformation. The shift from a Cartesian model of a subject to an intersubjective understanding of consciousness, subjects, and of the transcendental rules which govern them, paved the way to a re-conceptualization of reason in terms of a communicative model. And it is this new model of rationality that enabled Habermas to meet all the criteria that his ambitious attempt to answer the question of modernity had to preliminary meet.

Communicative reason, as Habermas conceives it, rests on the quasi-transcendent webs of linguistic norms. They are both part of the capacity to communicate and what in which subject and world primarily disclose and individuate themselves as self and world. These webs of norms provide a sort of Archimedean point which saves the unconditionality of reason without recurring to any metaphysical truth. Starting from the theoretical analysis of communicative reason, Habermas will shape a highly formalized system of procedural validation—Discourse Ethics—that, at least in his intention, will finally provide the needed foothold to reactivate the process of emancipation of modern reason. This potential reactivation, which Habermas sees as both pragmatic and theoretical, has tested itself, often recording a slow but persistent renegotiation of its equilibriums.

Habermas' refined model of rationality in general, and Discourse Ethics specifically have both undergone direct and substantial criticisms. My aim here is not to engage with his project in a similarly direct manner.[5] Instead, I intend to consider the efficacy of Habermas' endeavour to answer to the question of modernity looking at it from a sort of lateral perspective. That is to say, I investigate the functioning of discourse ethics from the perspective provided by those old fragments of past objectivity that one would have

[5] See for instance J. Tully, "To Think and Act Differently", in *Foucault contra Habermas*, Sage Publications, London, 1999; S. Benhabib, *Situating the Self*, Routledge, London, 1992, pp. 29-38; J. Thompson, "Universal Pragmatics", in *Habermas: Critical Debates*, MIT Press, Cambridge, 1982; R. Rorty, "The Contingency of a Liberal Community", in *Contingency, Irony, and Solidarity*, Cambridge University Press, Cambridge, 1989.

expected to be easily incorporated into Discourse Ethics, or to be demonstrated as now irrelevant. I am speaking of religion.

The issue of religion, instead, has, I will argue, indirectly, but constantly, highlighted the architectural inadequacy of Habermas' critical project.

The fragments of the old forms of objectivity that discourse ethics is supposed to appropriate and replace (or simply validate) have, over the years, proved more durable than Habermas originally expected. At the same time, discourse ethics has found itself in the position that it had to, in order to properly work, deal with specific semantic contents, which are already socially shared, and have in themselves a core of rationality that can be used as motivational substratum. This condition, that, at a certain point, Habermas himself defines as a "curious dependence", has pushed him to reconsider the role meaning plays within discourse ethics. In particular, he has engaged in a long-lasting and evolving challenge to adjust—while not resigning—his critical theory to those allegedly rational contents of the Judaeo-Christian tradition. A tradition that, in spite of its decline as both dominant worldview, and cognitive modality, not only did not disappear, but also proved to be particularly useful "for normalizing the intercourse" between discourse ethics and the very demand for meaning.[6]

Subsequently, and in a manner that intertwines, but does not identify with the evolution of the reflection about the mentioned religious and theological tradition, it became evident, first of all to Habermas, that the challenge that discourse ethics has to face does not exhaust itself in the research for rational allies. On the contrary, it expands from the cognitive level and the related functional problems which concern its

[6] Habermas, *Postmetaphysical Thinking,* The MIT Press, Cambridge, 1992, p. 51. From now on PMT.

implementation, to the meta-ethical sphere and the question about Discourse Ethics' conditions of existence.

More precisely, the neutral and universal functioning of discourse ethics is, in Habermas' view, shaken by the developing genetic engineering which menaces the integrity of the human species as known so far. The emergence of such a danger undermines, in Habermas' view, a form of unconditionality—the universal moral community of language users—which differs from the one provided by language, but that is as necessary as the latter when it comes to the very possibility of discourse ethics. This unprecedented danger pushed Habermas to look for a species ethics able to secure the existence of such a universal moral community, and, consequently, of discourse ethics.

On both mentioned fronts—the appropriation of religious contents, and the dealing with a meta-ethical dimension in the face of genetic engineering—Habermas' radically critical project proves to be unable to cope. My analysis of that project therefore focuses exactly on those two fronts, for I want to use them as test cases. And this because it is Habermas himself—explicitly and implicitly with the adjustments made to his work—who recognizes the relevance of those challenges. And their relevance is accentuated if I succeed in showing that Habermas' responses to these challenges are ultimately unconvincing.

With this chapter, after introducing central aspects of communicative reason and discourse ethics, I trace Habermas' attempt to deal with both challenges so that I can bring to light the manner in which he fails.

These failed attempts, the articulations of which vary according to the specific strategies adopted by Habermas at various stages, are, in my reading, the symptoms of a single problem. Discourse ethics, which is designed to be pragmatic in method, and

therefore *responsive* and open to facticity, shows itself *dependent* on specific features of that facticity. And the fact that these features have to be present or presupposed in order to secure the proper functioning of Habermas' version of radical critique implies that discourse ethics cannot be radically critical, for it cannot critically assess those very features that has to presuppose (as present). In other words, my argument is that discourse ethics presupposes specific elements in order to work, but these specific elements cannot be themselves validated by it. And given that radical critique requires the (at least potential) scrutiny of all the aims reason gives to or accepts for itself, discourse ethics fails to live up to that very demand.

Communicative Reason and Discourse Ethics

After an initial Marxist orientation which already linked emancipatory reason with communicative consensus and free dialogical relations, during the early 1980s, Habermas came to a theoretical structure that despite expansion and even some corrections, has remained constant since then.[7] Critical dialogues and disputes, in fact, along the years, have had the power to push Habermas to re-accommodate some of the equilibriums of the so called formal pragmatics, but what has been actually touched is not the very core of the theoretical construction; it has remained unchanged since that first mature articulation.

At the centre of Habermas interests is the normative foundation of a critical theory which is no longer entrusted to the alleged revolutionary power of the immanent contradictions of reality, neither is it left to a mere historico-genetic and in some way therapeutic reflection, rather it is secured by being linked with a "rational reconstruction"

[7] Concerning the very initial stages of Habermas' interest about the emanciparory sides of reason, see Habermas, *Knowledge and Human Interest*, Heinemann Educational, London, 1978.

of the “universal competences” present in every communicative act. The formal-pragmatic know-how enclosed into language, which is the ground even of the most basic communicative interactions, is employed to explain and regulate the intra- and inter-human relations. It is exactly in the structures that shape the possibility to communicate that Habermas indentifies the rational tie that enables the elaboration of criteria that can ground critical analysis without having to fear falling prey to an ideological perspectives or a nihilistic defeatism.

In 1981, after ten years of preparatory works, Habermas completed a theory of communicative action which embraces a decisive linguistic turn in his own critical theory. If the criteria of rationality of the single individual are unavoidably embedded in a formal and pragmatic structure which is intrinsically inter-subjective, then the primacy of consciousness cannot but leave space to the pre-eminence of the web of inter-subjective norms and meanings in and through which each subject discloses itself. It was the inauguration of a communicative model of reason in its most refined form that two years later brought about the complete formulation of Discourse Ethics.

Language subjects every speaker and every thinker to unavoidable norms, and shapes the rational basis of every human action. This means that in the structures of language is rooted a rational bind. This bind provides the criteria for a form of evaluation which justifies claims to objectivity, and therefore it can also be the source of a valid critical practice.

By reconstructing the normative frame present in language, Habermas discloses a cluster of validity norms and criteria for evaluation which, according to him, are not only the bones of procedures that can lead to consensus, but they are also the deep formal frame that permeate the shared common understanding of a community of speakers. It is,

in fact, the dialogically mediated interaction that allows individualization: for a reflexive autonomous identity to be an identity it must find itself already decentred into a web of norms which are, at the same, time medium and result of the individualization itself. In this way communicative reason is not only identity-shaping, but also the chance for the interacting subjects to mutually and consensually coordinate their linguistically mediated actions on the ground of the best available reasons; reasons that are sieved through a procedure in which they all already participate. The rightness of every communicative act is, in fact, determined by criteria of validity such as propositional truth (*Wahrheit*), normative legitimacy or rightness (*Richtigkeit*), and sincerity or truthfulness (*Wahrhaftigkeit*), that can be met or not, but that—even if still fallible—are a stable and already available foothold for practical rationality.

The reconstruction of the criteria of communicative rationality is conducted by Habermas through an analysis that is both phenomenological and analytical. Habermas rebuilds the quasi-transcendental condition of reason, and at the same time recognizes the dependence of such a form of rationality on the consciousness of the actual speaker, and therefore on the specific communicative actions and their institutionalized forms as occurring at a given time. In other words, his reconstruction must take into account the fact that the unconditional form of communication has its own time and realization: modernity. Formal pragmatic conditions of reason and the criteria of evaluation of its procedural functioning are thinkable only starting from a reciprocal presupposition.[8]

Furthermore, the so called linguistic turn situated as it is within a post-transcendental framework, allows a deeper level of analysis. If the transcendental point of view at the base of the Kantian edifice instead of completely disappearing has migrated

[8] TCA I, p. 16.

from the ontological core of subjective consciousness to the normative and constitutive webs of language, what follows is that the limit that separates the subjective from the objective becomes porous, while the two elements become reciprocally receptive. The decentred quasi-transcendental webs of language not only are irrefutable *and* historically placed, they also play an essential role in the formation of identity. The individual, in fact has the chance to recognize both herself and the world only through the opening of an alterity which is never a radical other. This sort of co-belonging makes the individual originally intersubjective.[9]

A final consequence of the dissolution of a transcendental *I* is that reason is conceived as completely situated. In a post-metaphysical time, in which language that makes the world accessible to us also shapes and makes possible our relation to our self through the mediation of "the subjectifying gaze of the other", reason itself and all its articulations (communicative, instrumental, and aesthetic) find themselves situated within the historical flux, the grammar of which is already and by necessity dictated by language.[10]

In this way communicative reason and its historical-theoretical correlate place themselves at the centre of a theory of modernity that aims not just at self-comprehension, but also at the advancement of rationality towards a progressive autonomous status. The aim, for Habermas, is to balance plurality and unconditionality within the practical field in order to ensure an egalitarian humanism. A process understood not as mere rationalization of the systemic spheres of society—as economy and bureaucracy, which, according to Habermas, are forms of rationality independent from, and often colliding with, the communicative sphere—but above all, as rationalization of the cognitive, normative, and expressive elements of daily life.

[9] Habermas, *Between Naturalism and Religion*, Cambridge, Polity, 2008, p. 16.
[10] Ibid., p. 15.

Communicative reason provides, according to Habermas, both the normative criteria necessary for a non-certain evolutionary logic of social and cultural formations, and an already shared unconditional standpoint. Therefore it provides the essential pre-condition for a non-destructive critical theory that regards facticity as a sort of material that needs to be tamed and employed.

From the theoretical point of view, the most important consequence of the theory of communicative reason is that it allows for the construction of a Discourse Ethics that is able, at least that is Habermas' intention, to bring about an intelligible and functional distinction between the whole historical determined ethical sphere, and morality. That is to say, discourse ethics, as rational procedure of practical argumentation, must be able to select from the entire cluster of ethical life—which encompasses interpersonal relations, institutions, habits and traditions—only a set of universally valid norms, that is to say, morality in Habermas' understanding. The validity of those norms is provided exactly by the formality of the procedure—discourse ethics—through which they are selected.[11] In this way, paraphrasing Habermas, the moral norms are differentiated and therefore insulated from questions of a good life—precisely the ethical questions which do not lend themselves to a direct and feasible universalization.[12] This means that the contents, and the cognitive modality that constitute a good life are considered as possible candidates to be turned into norms, yet they are not norms, unless they pass the scrutiny of discourse ethics.

This theory of morality, which is developed against the background of an evolutionary vision of rationality, is, in Habermas' work really detailed. First of all, what Habermas stresses is that discourse ethics as a philosophical theory, and therefore as reflective thematization of moral norms, is not a mere description of the contents

[11] Habermas, *Moral Consciousness and Communicative Action*, Cambridge, Polity Press, 1992.
[12] Ibid., p. 108.

expressed by those norms, but it is a formal function aimed to justify *the moral point of view.* And the way this moral point of view can be justified, given the communicative model of reason, is through an argumentative dialogue.

Such a dialogue, if it is to produce valid norms, must follow two principles. It must follow the *Principle of Universalization* [U] according to which "*All* affected can accept the consequences and the side effects its general observance can be anticipated to have for the satisfaction for *everyone'*s interests", and, at the same time, the *Discourse Principle* [D] must find application too.[13] That is to say, "Only those norms can claim to be valid that meet (or could meet) with the approval of all affected in their capacity *as participants in a practical discourse.*"[14] While [U] already refers to the necessity of consensus but mainly focuses on the necessary condition that makes that kind of consensus universal, [D] expresses the essentiality of argumentative consensus in order for something to be considered valid. In other words, the two principles aim to bring together the two sides of the universality: universal access and universal validity.

Such a procedure, once it has reached its most refined form, finds itself unable to work in the very socio-historical context which yet caused the emersion of its necessity.

The steering processes that are trigged by the collapse of the old order – which is seen by Habermas as an essential passage to the liberated and liberating forces of communicative reason – present themselves as a difficult obstacle to overcome also within this critical model. Processes of rationalization, as Habermas the sociologist recognizes, have weakened the social fabric and consequently the essential motivation which would push people to unleash those already available communicative forces in order to reactivate the spirit of modernity.

[13] Ibid., p. 65.
[14] Ibid., p. 66.

The lack of shared socio-political motivations—"as utopian oases dry up, a desert of banality and bewilderment spreads"—together with the inability to create alternative contents encompassing sources of rationality, led Habermas, against his initial expectation, to deal with a 'moment' of reason that along his work has shown itself not inclined to be reduced to the status of "contemporaneity of the non-contemporaneous".[15]

Habermas faces two opposite, but equally dangerous, possibilities for communicative reason: either to shrink to a mute "grandiose tautology" or to open itself up to a nullifying Babel of ethical relativisms. Against this backdrop he seeks a possible alliance between a secularized practical reason and religion.

The search for the proper place for religious contents within a modern form of radical critique is, in Habermas' work, a long, evolving, and ultimately unsuccessful process. Trying to reduce this process to an essential numbers of stages, I will refer to three separated and consecutive moments: sublation, coexistence, and cooperation.

Sublation

As the outline of Horkheimer's analysis has already indicated, modernity's process of rationalization puts its critical spirit as well as reason itself in a great danger exactly through its own logic. Habermas, since the beginning of his intellectual itinerary, has shared this analysis, but only in order to look for a way to avoid this impasse of the *Eclipse of Reason*. It is exactly the search for the practical and theoretical conditions of the possibility to overcome this impasse that brought him immediately into contact, even if just superficially, with the concept of religion.

[15] Habermas, *The New Conservatism: Cultural Criticism and the Historian's Debate,* Cambridge, Polity Press, 1991, p.68.

For Habermas, the necessity to both have a clear understanding of the manifold problems which have led the first generation of the Frankfurt School in a sort of deadlock, and to look for a possible space of theoretical manoeuvring, has put him in the condition to have to retrace first in sociological, and then in philosophical terms, the *status quo* of the modern malaises.

In particular, if an efficacious answer had to be found within a communicative model of reason, then meticulous attention had to be paid to the state of the semantic sources which permeate, and shape the community of speakers. The vital processes of production of meaning, and its distortions need to be traced and opened to future therapeutic interventions. And it is within this kind of preliminary study that religion, as one of the most important past sources of meaning, is considered by Habermas.

During 1973, Habermas wrote a tentative and open analysis of the practical difficulties of the Marxian theory of crisis of the advanced capitalist societies in which he simply aimed to provide a systematic and integrated account of the state-of-the-art of the most advanced studies on the matter.

On that occasion, he describes the role religion has as practically inexistent. The key term is '*sublation*'.[16] Within a conception of rationality still strongly depended on the Hegelian-Marxist dialectic, Habermas looks at religion as an overcome stage of the process of rationalization which is no longer relevant, not even within the private sphere, if understood as cognitive modality.

At that time, the evolutionary process of both reason and society appeared to Habermas, at least with respect to religion, straightforward: religion as meaningful frame within which to make sense of reality is something belonging to the past. Yet, its contents,

[16] Habermas, *Legitimation Crisis*, Heinemann Educational Book, London, 1976, pp. 79-81.

or at least some of the contents of those grand narrations, have been incorporated by the modern processes of rationalization. The transformed secularised utopian messages of salvation, for instance, which belong to the Judaeo-Christian tradition, and that still remains active in modern societies, are described by Habermas as first unfastened from a general order of meaning, and then by the means of that unhooking, exposed to the rapacious exploitation carried on by what he defines as the subsystems.[17]

Habermas explains that once a concept like salvation, or whatever other semantic content, is no longer part of a precise horizon of meaning which sustains it by providing it with a general frame of intelligibility, then that fragment is exposed. And it is exposed in a double manner. First, it needs to face the competition of alterative views, the mere existence of which call into question its validity. Then, the crisis of its validity leaves it open to possible exploitation. For once it is no longer universally valid, it can be employed in the service of a one-sided political view, of some institutional policy, or of a commercial interest, for instance. These kinds of instrumental use will wear-out the original meaning, thereby weakening, sometimes in an irreversible manner, its force.

The exploiting processes that subsystems carry out is for such subsystems vital; for as self-reproducing and formally integrated domains mediated by money and power, they do not have the capacity to produce the content they employ.

This, in itself, would not be an insurmountable problem if modern societies were able to produce new semantic sources capable to substitute the exhausted ones. Though, to the processes of exploitation performed by the subsystems, must be added the incapacity of the modern societies to generate new semantic contents. As a consequence,

[17] Habermas borrows the concept of 'system' from Parsons, and Luhmann. See Parson, *The Social System*, Routledge & Kegan Paul, London, 1951, and the joint work Luhmann/Habermas, *Theory of Society or Social Technology. What does Systems Research* Achieve? published in Frankfurt in 1971, [J. Habermas/N. Luhmann, *Theorie der Gesellschaft oder Sozialtechnologie – Was leistet die Systemforschung?*, Suhrkamp Verlag, Frankfurt am Main, 1971].

what Habermas defines as “motivational syndrome” and “syndromes of civil and familial-vocational privatism amounts to a recurring crisis which hits both societies and individuals. Furthermore, it is extremely difficult to imagine a strategy that is able to interrupt this vicious circle.[18]

In this context, despite the description of this detrimental logic concerning the state of meaning, religion is, for Habermas, not yet the name of an issue, let alone a problematic one. It is simply part of a genetic analysis of semantic contents, and of the processes of socialization which Habermas uses to illustrate how and why modern societies have reached such a state of deep crisis. Semantic contents, such as the secularized components of religion, are scrutinized in order to show the role they play in both the constitution and preservation of a social fabric, which risks to be weakened to the point of eliminating the potential for non-technological progress.

At this point of the evolution of Habermas’ thought, there is no clear and worked-out positive critical theory that could deal with or avoid that risk, and the domino effects that are triggered by it, but he has already shaped some fundamental theoretical instruments that are employed in the considered work as grid to both interconnect the different levels of analysis of the modern crises, and spot at least the conditions for a possible way out.

In order to understand (a) why semantic contents as such were not yet a problem in themselves, (b) why they have become so, and (c) what was the state of Habermas’ answer to the question of modernity, it is worth engaging a bit closer with what he meant by stating that tradition no longer exists in its original shape, and yet it is appropriated and dissipated by the highly formalized functions of modern reason.

[18] Habermas, *Legitimation Crisis*, Heinemann Educational Book, London, 1976, pp. 75-77.

Once the processes of rationalization of modernity have started, a precise dynamic has been established according to the philosophical tradition on which Habermas is relying in a non-orthodox way.[19] The reproduction of the merely functional elements of reason, the grammar of which shape labour, the political, and administrative organization of the state, is something that even if increasingly autonomous is not autarchic.

These subsystems, in fact, cannot generate the preconditions for their reproduction. They need to be fed with contents and behaviours that pertain to the relational sphere of society, that is to say, the one that directly obeys to the communicative level of rationality as opposed to the functional one. Yet, this productive and reproductive relation is systematically unbalanced, and therefore bearer of structural problems, for, according to Habermas, the relation between the relational elements of reason and the functional ones is both practical and ontological. They are both shaped and kept alive within processes of socialization which, in turn, take place within structures of linguistic intersubjectivity. And it is within this intersubjective media (language, culture, and myth) that every possible action—purposive-rational action included—and its motives find their condition of possibility.

Given that this condition is essentially linguistic, it is originally bound by validity claims; otherwise, as Habermas notices, that form of cultural life would not be able to reproduce and survive. This means that the validity of symbols, narrations, habits, and beliefs, which have contributed to shape the intersubjective consciousness of the individuals, must be potentially redeemable through argumentation/consensus.[20]

[19] Habermas thinks that increasing the rationality of the subsystems which work by and through the production of crisis that can bring society to the threshold of anomie, is not the same as increasing the rationality of the communicative sphere. This distinction, which is clear since *Knowledge and Human Interest,* provides a way to discriminate between modes of rationalization and is probably what marks the moment of Habermas' major distance from the first generation of the Frankfurt School.

[20] Habermas, *Legitimation Crisis*, Heinemann Educational Book, London, 1976, pp. 5-6.

Contrary to the relational processes, the instrumental elements, despite them sharing the originally linguistic root, work *de facto* without a direct justification. They work by relying on technical rules which need cognitive and semantic sources no longer to orientate their actions, but merely to keep themselves alive against and through the material they work with.

According to Habermas, these two paradigms affect and hinder each other to the detriment of the entire process of modern emancipation.

The media of instrumental reason (money, bureaucracy, and socio-political institutions) need to appropriate, in order to work, cognitive contents coming from the communicative sphere as "there is no administrative production of meaning".[21] This kind of appropriation is made possible because the original source of the meaningful contents is no longer valid, and therefore it has no force to oppose the appropriation.

What is more, the sources of meaning available are not only unstable since they lack a valid framework, but they are also unable to renew themselves. Bourgeois culture, according to Habermas, has never been able to produce and or renew itself "as a whole". It, in fact, has been relying on semantic sources that are not its own. And since the pre-modern order that justified them is no longer valid, bourgeois culture is compelled to use those sources as if they were its own semantic contents. Such contents, which are used as "motivationally effective supplementation", are, however, a weak solution, as bourgeois cultures are unable to provide any useful narration aimed to bridge the communicative needs and an objectified nature that multiplies the contingent "risks of existence". The

[21] Ibid., p. 93.

consequence is that the semantic exploitation realized by the subsystems to the detriment of the worldviews endangers the society as a whole.[22]

In particular, once the practice of solidarity, which is understood by Habermas as the fundamental pillar of the social fabric, is formalized by state institutions, it is no longer easily available to society which had originally cultivated and embraced that solidaric practice. The formalization and abstraction of solidarity, instead of improving the force the original motivation, result in what Habermas defines as "steering problems", that is, mere automations.[23]

The transformation of the practice of solidarity into an increasingly formal function leads solidarity to be alienated from its sources. Once this happens, its automatic function reduces almost to zero the political participation of the citizens. This, in turn, has a double effect: the semantic impoverishment of the social fabric, and its distortion. For, once semantic contents are both deprived of their original validity and then impoverished by the subsystems, the self-comprehension of the citizens is, at very least, hindered. In this manner, it is the totality of the processes of expression that are endangered.[24]

It is for this reason that the most worrying consequence of the erosion of meaning is that "the intuitive access to relations of solidarity within groups or between individuals" is made impossible. This means that the motivational condition of the possibility for a renewed political and ethical dimension of society becomes unavailable. At least this is the scenario that Habermas sees when he turns his analysis on the state of the semantic sources.[25]

[22] Ibid., pp. 78-79.
[23] Ibid., p. 4 et passim.
[24] Ibid., p. 110.
[25] Ibid., pp. 78-79.

The "cognitive dissonances" between what is needed and what the subsystems automatically impose and distort manifest themselves in three main ways: a) the world-views lose the ability and even the ambition to provide their own account of the reality as whole; b) religious beliefs which were already relegated to the private sphere – without finding any room or role in the shared social practices – are now relativized once more, transforming from a private understanding of reality into fragments of a vast competing plurality of beliefs "that is undecided as to truth"; c) moral conceptions, detached from the holistic systems of interpretations, leave space only for an "utilitarian secular ethic" within which egoism is not only seen as unproblematic but rises to the level of "common sense".[26]

Being stuck in an ethical-political stasis that does not allow for any successful substitution of the old semantic sources with new or simply renewed ones that are able to provide a direction and motivation, it is licit to wonder if there is an actual practicable solution to this impasse. It is Habermas himself, going back to the question that pushed him to engage with the analysis of the modern crises, who wonders:

> If world-views have foundered on the separation of cognitive from socially integrative components, if world maintaining interpretive systems today belong irretrievably to the past, then what fulfills the moral-practical task of constituting ego- and group-identity? Could a universalistic linguistic ethics no longer connected to cognitive interpretations of nature and society *a)* adequately stabilize itself, and *b)* structurally secure the identities of individuals and collectives in the framework of a world society?[27]

[26] Ibid., pp. 80-81.
[27] Ibid., p. 120.

And the question is a real one. For, in a postmetaphysical time and given the described context, the possibility that complex societies simply allow for no normatively justified practical motive formation must not be underestimated. Yet, according to Habermas, the possible, and indeed necessary replacement will happen due to a communicative ethics that—forfeiting the direct engagement with meaning—will prove itself reliable but not comprehensive, for structurally securing the identity of the social actors, and guaranteeing their societies' universal framework.

The aim, still haunted by the doubt of a possible failure, is to identify the pragmatic-theoretical conditions for a universal morality that is able, without an ethical root, to not "shrink to a grandiose tautology", that is, to a sort of moral mutism.[28]

Even without being able to provide solid arguments in support of the fact that the universal morality will find space within the framework of modernity, Habermas seems to find a hint of a foothold against moral pessimism. He states that thanks to "the re-politicization of the biblical inheritance" it is possible not to liquidate God—which is here to be understood as synonymous with an unconditioned point of view—but rather to transform that idea into a dialogical logos that is able to reactivate both the identities of the individuals and the concept of a "self-emancipating society". This proto idea of a discourse ethics is at that time more a suggestive image than a theory:

> The idea of God is transformed [*aufgehoben*] into the concept of *Logos* that determines the community of believers and the real life-context of a self-emancipating society. "God" becomes the name of a communicative structure that forces men, on pain of a loss of their humanity, to go beyond their accidental, empirical nature to encounter one another *indirectly,* that is across an objective something that they themselves are not.[29]

[28] Ibid., pp. 120-121.
[29] Ibid.

The pathological restriction of communication caused by the formalized elements of society during the 1970s are seen by Habermas as a distortion that, under the threat of the most radical of the loss, might be overcome by a form of unconditionality which aims to replace God.

God, as what must replace it, is not conceived as a cluster of beliefs, but rather as a sort of empty agglutinating structure that could uncoercively force each individual, through a medium, to recognize their co-belonging as essential to the very possibility of a morality which aspires to be unconditioned.

While God as the general framework to replace finds space in Habermas' view of his theory to come, religious beliefs, and their secularized narrations, on the contrary, are simply considered as possible sources of inspiration. The "biblical inheritance", whose original meaning is understood by Habermas as having no longer currency, plays, at this stage, no role in the formation of a new moral perspective.

At this stage of Habermas' analysis, the contents of religion are seen as digested by history and in no real need for revival, while the role of the (unconditioned) ordering framework which lays the conditions of what can be considered true is in need of replacement since its position is vacant.

The absence of what will later be discourse ethics is, according to my reading, what makes religious contents unproblematic. Things will gradually change in Habermas' view, and the case of religious content will be enlightening about the reasons that propel the very change.

Coexistence

Both the declaration of the death of the religious contents, and the anticipatory vision of a progressive and definitive overcoming of the role of God had to be revised as soon as Habermas' projects for a universal pragmatics started to be refined. Yet, the revision of his previous theoretical position, and the consequent passage to his new understanding of the relation between his pragmatics and religion, occurred in a gradual and non straightforward manner.

Starting from the 80s, Habermas retrospectively recognizes his own debts to the religious tradition, and in a particular way to the Jewish-Christian one, as main historical source of a universalistic ethics. In the fundamental works of those years, he looks at the mythical, religious, and metaphysical forms of thought as essential genetic stages of the rational evolution (learning process) of the structural elements of communicative rationality; though, he denies that the contemporary stage of practical reason still needs any of those forms in order to be effective.[30] Starting from the Enlightenment, in fact, practical reason has made itself autonomous from its origin, not only distancing itself, but trying to replace it. So much so that Habermas writes:

> With the transition to a new stage the interpretations of the superseded stage are, no matter what their content, categorically devalued.[31]

[30] For the whole phylogenetic account of the evolution and emancipation of the modern normativity as linguistification of the sacred, see TCA II, pp. 43-111.
[31] TCA I, p. 68.

Religion is in this period part of a phylogenetic reconstructive account of the forms of rationality, which from its very basic magical-animistic stage has arrived at the secular one, that is to say a specialized, diversified and reflexive version of reason whose universalistic form substitutes the explicative power previously associated with religious narrations.[32]

This means that communicative reason in its first stable configuration confirms the hypothesis present in *Legitimation Crisis*, that is to say, Habermas understands it as a sublated stage of the idea of God.[33] Yet, at same time, he does not re-state his position about both religious contents, and the idea of God in exactly the same manner. It is, in fact, possible to see that, even if the idea of an Hegelian negation of the moment of religion remains unchanged until the second half of the 80s, the way in which Habermas deals with this issue starts to evolve contextually to the formalization of the concepts of communicative reason in the 1981 and of discourse ethics in the 1983. For instance, resolute statements as strong as the ones in *Political-philosophical Profiles* about the irrelevance of any discourse around God disappear from Habermas' work before the actual theoretical change.[34]

Retrospectively one can see that those rhetorical aspects of Habermas' writing mirror the emerging of a new theoretical concern. It is, in fact, during the years in which Habermas was engaged with the shaping of a communicative model of reason, that he must have become more and more aware of the necessity to look for a theoretical equilibrium capable to go beyond the dual and unproductive dispute between God and its death. And it is the very search of such an equilibrium that first showed the signs of a

[32] See all last chapter of TCA I and TCA II: "the sacred domain has largely disintegrated, or at least has lost its structure forming significance", p. 196.
[33] TCA I, see in particular pp. 67 and 164, and the extensive use of Weber and his understanding of 'sublimation' of religion by knowledge.
[34] Habermas, *Political-philosophical Profiles,* MIT Press, Cambridge, 1983, pp. 1-20.

change *in fieri*, and then contributed to Habermas' recognition that the process of sublation is, at least, not completed:

> Communicative reason does not make its appearance in an aestheticized theory as the colorless negative of a religion that provides consolation. It neither announces the absence of consolation in a world forsaken by God, nor does it take it upon itself to provide any consolation.[35]

1988 is the year in which there is a first substantial and explicit change. With the publication of *Post-Metaphysical Thought* Habermas, after having realized the wrongness of the equation: disenchanted society equals atheistic society, exhibits the results of such changes. These changes concern both his analysis of religious forms of reasoning and the mere vitality of religious believes that, according to his previous evolutionary account, were on their way towards extinction.

Habermas now recognizes without any indecision the coexistence, legitimacy, and indispensability of religion, going as far as to wish for the survival of religion besides the theoretical and practical modern form of reason, at least until the moment at which reason alone will be able to fill the motivational and cognitive deficit of modern societies.[36] While the recognition of the "ongoing coexistence" is stated without any difficulty or ambiguity, Habermas is vague on the status of such a coexistence.[37] It is not completely clear whether it is a temporary coexistence or not. What is more, if there is coexistence, does that imply

[35] PMT, p. 145.

[36] PMT, p. 51. About this point see P. Dews, *The Limits of Disenchantment*, Verso, London, 1995, p. 10.

[37] Hermann Düringer, as N. Adams (2006, 12) notices in his monograph about Habermas and theology, suggests a similar reading of the relation between Habermas' understanding of philosophy and religion. Düringer refers to the relation as "unmixed" and "unseparated". That is to say, according to Düringer, Habermas, at that stage, sees religion as unrelated to the practice of post-metaphysical philosophy, yet he recognizes that the German tradition on which his own work draws has mystical roots. H. Düringer, *Universale Vernunft Und Partikularer Glaube: Eine Theologische Auswertung Des Werkes Von Jürgen Habermas*, Leuven, Peeters, 1999, pp. 19-25. [Universal Reason and Particular Faith: A Theological Evaluation of the Work of Jürgen Habermas]

some form of dependence and/or collaboration? Habermas' theoretical position at this point is not easy to decipher as every statement that he makes about this specific issue is mitigated by the use of prudential expressions.

In two different key moments Habermas maintains that philosophy will replace religion. Yet this is the way the two passages read: (i) "[philosophy] will be able neither to replace nor to repress religion as long as religious language is the bearer of a semantic content that is inspiring and even indispensable, for this content eludes *(for the time being?)* the explanatory force of philosophical language and continues to resist translation into reasoning discourses."; and then (ii) "As long as no better words for what religion can say are found in the medium of rational discourse, it will even coexist *abstemiously* with the former, neither supporting it nor combatting it".[38]

In the first case it is clear that Habermas does not want to expose himself to a new unfortunate prediction, while the second is more problematic because that 'abstemiously' contrasts with another assertion that he makes in the same text and that escapes a simple interpretation; that is to say the definition of the relation between religion and philosophy as "a curious dependence" of the latter on the former.[39]

However one is inclined to interpret those quotations, what is clear is that Habermas has changed his view. And this change is so fundamental that it is Habermas himself who emphasizes it. Furthermore, he makes it clear that this change has to be understood as forming part of the wider attempt to critically react against the endeavour to renew, in the wake of German idealism, metaphysical thought. At least this is how Habermas presents the evolution of his own work.[40]

[38] The emphasis in both quotations is mine. They are respectively taken from pp. 51 and 145 of PMT
[39] PMT, p. 51.
[40] Ibid., pp. 14-18.

Philosophers such as Dieter Henrich, who reckons that there is a necessary connection between political and philosophical theory and the Absolute and with the fundamental question of the humanity of the humans, are presenting, according to Habermas, theories which belong to a field that is not the philosophical one.[41] They, according to Habermas, would be more appropriate within a religious stance, rather than as part of any critical rationality that deals with knowledge and that as such wants to preserve its validity by acknowledging its own limits.

Yet the engagement with religion and meaning cannot be explained only on the grounds of a simple reaction to allegedly dangerous neo-metaphysical theories in order to make sure the fields of competences do not overlap—after all, it is Habermas himself who declares his engagement with religion in a manner that goes far beyond its use as shield against metaphysics. The real issue is that religious languages, even if dis-empowered of their ability to provide an encompassing account of the world ("its word-view function"), can no longer be treated like a "contemporaneity of noncontemporaneity".[42] And this is because philosophy—and by philosophy Habermas means discourse ethics—finds itself caught in the "curious dependence" I previously mentioned.

The quotations come from a chapter called *A Return to Metaphysics?.* In this chapter, Habermas argues that philosophy is unable to deal with what is defined as "the subversive intrusion of the extraordinary". This inability that is described by Habermas as a characteristic feature of a post-metaphysical critical thought, demands an external support for "normalizing" precisely the relation between a discourse with claims of validity, on the

[41] Henrich, D. 'Self-consciousness: A Critical Introduction to a Theory', in *Man and World,* vol. 4, 1971.
[42] PMT, p. 51.

one hand, and experiences that involve forms of meaning that can be hardly validated, on the other.[43]

Religion, with its cluster of semantic sources, is seen at that point as a beneficial aid. It, in a way that Habermas does not fully explain, supports, and even inspires people in their search for meaning. What "extraordinary" stands for is not clear, Habermas does not make it explicit, but probably he is thinking about all those contingent instances of the negative, such as death, for example, which forms part of everybody's life.

If the meaning of the "extraordinary" is obscure but not fundamentally problematic, because one can easily assume that a post-metaphysical philosophy has to withdraw its limits within what can be rationally appropriated and validated; the same cannot be said about what Habermas meant with the "better words" that religion would be able to provide, or with its "semantic content" that would be "even indispensable". Habermas, in fact, holds that religious semantic resources remain not only alien to a formalized reason but, at the moment, they are even resistant to any rational attempt to render them into a reflexive form.

Now, if the contents and concepts towards which Habermas is gesturing are the emancipatory ones ingrained in the religious messages of salvation, then it is difficult to understand why a critical discourse would be unable to appropriate them, since it is Habermas himself who writes:

> I do not believe that we, as Europeans, can seriously understand concepts like morality and ethical life, person and individuality, or freedom and

[43] Ibid.

> emancipation, without appropriating the substance of the Judeo-Christian understanding of history in terms of salvation. [44]

According to this statement, those contents are not only products of the same historical process that has produced the conditions of possibility for discourse ethics, but they are also the "semantic potential" that every new generation needs to appropriate again and again in order to "be able to recognize him- or herself in all that wears a human face".[45] This recognition of each individual as one's equal is thus part of a learning process that precedes discourse ethics both logically and historically. Therefore, such substance, according to the genetic account that Habermas himself provides, is already available to the reflexive processes of self-understanding. This automatically excludes the inability for discourse ethics to appropriate the idea of a historical process understood as a form of possible salvation; and it also leaves, at least, two unanswered questions: what are then these semantic contents that rational discourse is unable to appropriate? And b) why, and on what ground, are they deemed as better compared with the semantic contents discourse ethics can actually appropriate?

It is, in fact, at least "curious" that Habermas, while asserting the inaccessibility of those sources, does not avoid evaluations concerning the positive role they play.

As the questions remain unanswered, it is possible to try and infer a plausible explanation.

It seems that Habermas individuates in those alive religious languages a semantic surplus which is not present in the secularized religious concepts that discourse ethics can already use. Such a surplus, even if not clearly identified, has the peculiar power both to

[44] PMT, p. 15
[45] Ibid.

resist the erosion practised by the autonomous subsystems, and compensate, in a non elitist manner, the space necessarily left empty by the withdrawal of the rational discourse within the limits of what can be validated.

This interpretation can be supported by recalling the repeated use of the idea of consolation, which Habermas associates exclusively with religious languages. This concept is understood by the Habermas of those years, as something different from a secularized understanding of salvation—probably because "consolation" has a more specifically experiential connotation, if compared to salvation. And the contents of that experience—whatever they are—are, at the same time, more likely to have a stronger motivational grip, and less likely to be turned into self-reflective knowledge.

However, if Habermas is unable to explain what that semantic surplus is and why, even if the constant action of communicative reason is unable to grasp it, we should value it, then it is really unlikely that he can legitimately claim the proper functioning of discourse ethics since it relies on the support of something that cannot be assessed, not yet, at least.

Furthermore, if as Habermas states in that same chapter, the "syndrome of validity" on which religion relied on has been "dissolved by the emergence of expert cultures in science", then their temporary necessity as a specific form of self- and world-comprehension, remains unexplained not only form the point of view of secular reason, but also from the point of view of those experiences of the extraordinary which call for comprehension, but that have validity sources nobody can access or recognize rationally.[46]

It seems evident that Habermas was, at that time, dealing with some deep problems of discourse ethics, which in 1988 still have not found a clear articulation. The semantic and expressive powerlessness of discourse ethics had the role to rationally ground the

[46] PMT, p. 17.

appropriation of meaning. Its normative function had to remain empty because it was supposed to receive, evaluate, and finally give back to moderns legitimated contents. This formal process included the chance to use also the expressive elements of secularized reason; instead at this point of Habermas' work, discourse ethics presents itself against a background of a non-relational coexistence with meaningful narrations which shows nothing more than a methodological inconceivability. We can see, on one side, Habermas the sociologist who recognizes both the fact of religion, and a sort of quest for meaning, and, on the other side, Habermas the philosopher who has to accommodate anew the equilibriums of a formal procedure which has difficulty both in activating itself, and in dealing with the semantic sources of society. Habermas' difficult task of bridging the sociological dimension with the philosophical one appears, at this stage, doomed to fail. Religious contents are both needed, and inaccessible. Thus discourse ethics finds itself in a state of aporia.

However, Habermas' obscure account of this sort of unbridgeable proximity between discourse ethics and semantic sources is probably more an exercise to understand what discourse ethics is capable of rather than shaping anew his answer to the question of modernity within the historical and theoretical context that made that answer urgent.

In order to understand that, Habermas makes in *Between Metaphysics and the Critique of Reason* an attempt at clarifying what discourse ethics cannot do. It cannot provide any consolation or any explanation that would harmonize cognitive needs and contingency, and even in its most effective possible version it cannot remove that conflict, for contingency can be neither ignored nor fully explained. In this regard, at least at this

stage of rationalization "no better words for what religion can say are found in the medium of rational discourse".[47]

Furthermore, and this is just an elaboration of the previous point, discourse ethics cannot provide a substitute for any Absolute. Hence, the idea of a rational function able to substitute the role of the idea of God, as suggested in *Legitimation Crisis* has been, at least, modified.

The form of humanism Habermas battles for with his discourse ethics has now the shape neither of a self-asserting humanity, nor of its negation. At this stage, religion, as form of private belief is still alive, while God—the idea that legitimates religious contents as a form of knowledge—is dead and cannot survive in the shape of a sort of theology of death. For, according to Habermas, even that form of negative theology would be a way for metaphysics to survive in its negative: a speechless point of view hypostatized into the negation of all of what metaphysics once stated. In other words, the definition of discourse ethics as the negation and the overcoming of the function of God would risk saving a metaphysical fragment. For a definition based on a negation still takes the negated thing as its model. While according to Habermas the only fragment of simile-metaphysics that modernity can allow for is the unconditionality "preserved in the discursive concepts of a fallibilistic truth"[48]. And this unconditionality, which, for Habermas, is the foundation of his understanding of morality, "is not an absolute, or it is at most an absolute that become[s] fluid as a critical procedure"[49]. Such a procedure is unconditional due to the fact that its working is characterized entirely from a formal point of view. It is exactly its formality that can provide fallible claims to truth.

[47] Ibid., p. 145.
[48] Ibid., p. 144
[49] Ibid.

In addition to the intrinsic limits of a formal procedure, there is another fundamental thing discourse ethics cannot do. A formal procedure has no power to bring about the only thing that would make it work: a combined effort "marked by solidarity".[50]

Once it is clarified what discourse ethics cannot do, Habermas finds himself with a functioning unconditional validating procedure without motivation on one side, and, on the other, motivations—religious ones *in primis*—with no propensity to be validated. It is at this point that Habermas starts to prepare a typical Hegelian move that will introduce a concept of translation. It will become apparent that this amounts to nothing more than the attempt for an (im)possible recognition.

Cooperation

The secular *Diskurs* has no words to efficaciously compete against languages able to provide meaningful answers to the experience of negative and support to and consolation for mortal speaking beings. In addition, the creation of those words is something which is completely precluded to a rational moral procedure. Concepts such as salvation and consolation or any other way of dealing with the "extraordinary" (Habermas' umbrella term under which he groups together all the emergences of negative elements of contingency within the rational fabric of reality), which until 1988 were considered obsolete, have forced themselves to the threshold of discourse ethics without having a real chance to permeate through a logic that remains alien to it.

What is striking about the collection of essays published in that year is that Habermas goes on building his thematization of the coexistence of communicative reason

[50] Ibid., p. 146.

and religion without actually engaging with the meaning of those semantic sources. He exclusively appeals to a sort of ancestral hope for salvation that philosophy as procedure obviously cannot fulfil without violating its own condition of post-metaphysical validity. A hope that, on the contrary, religion addresses, fulfils, and makes intelligible by providing narrations whose contents are doubly useful, as Habermas clearly sees. On the one hand, the concept of universal salvation that comes from a secular elaboration of the Judeo-Christian tradition works as motivation for engaging in discourse ethics; on the other, the same tradition provides (or could provide) the pieces of rationality, such as the idea of equal dignity, that discourse ethics must presuppose in order to work properly.

Having opted for a normative procedure, Habermas cannot but renounce to any cognitive-ethical interference given his aim to preserve a possible universality. Discourse ethics cannot be more than a decentred normative standpoint. This means that while its proper functioning is not granted, in order to have a chance, it needs to presuppose some conditions. Discourse ethics requires that each potentially moral agent, in order to adopt the necessary attitude, may understand him/herself and the others like members of the potential moral community. This, in turn, requires semantic and motivational sources that communicative reason presupposes but which cannot be part of its procedure.

How these semantic and motivational sources may possibly be activated without a global meta-ethical perspective remains an unanswered question, or better an eluded one, hidden behind the reassuring switch ethics/religion. The attention dedicated to the issue of religion, in fact, fuels the suspicion that, at least until a certain moment, it has been easier for Habermas to deal with beliefs than to deal with the ethical dimension. Beliefs *qua* beliefs have no universal claim of validly, or at least that validity is relegated to the private sphere. Ethics, or even a weak meta-ethical shell, on the contrary, have claims of validity

that are both needed and incompatible with discourse ethics' formal validity as I will show later in the section on the ethic of the species.

This difference makes Habermas look at religious contents as both less problematic and, in a way, more useful than ethics given the structure of discourse ethics. Particularly, if it were possible to find a way to appropriate only what is useful to communicative reason without falling prey to a cognitive attitude that can no longer be considered to be congenial to philosophy.

It is towards this direction that Habermas' work moves. Though, in the years in which Habermas envisioned a mere coexistence without any form of collaboration between discourse ethics and religious contents, the appropriation of those contents appeared as blocked by what, in the previous section, I have defined as an unbridgeable proximity.

Considered within the experiential dimension, religious contents are not prone to be appropriated and consequently used within any argumentative function. For philosophical reflection alone cannot penetrate the experiential as such, especially when those experiences belong to someone who, relying on revelation, may have no interest in articulating those very religious experiences in the form of augmentative thought. And yet religious experience, according to Habermas, is one of those elements that better than other can help us "to keep this sense of humanity alive and to clarify it".[51]

The search for a possible critical appropriation of the substance of those experiences will lead Habermas to start dialogues with both theologians and philosophers who, in an effort of self-understanding, try to thematize religious contents from within the

[51] PMT, p. 15.

religious sphere, providing, in this manner, semantic elements accessible also to those who do not share the religious experience itself.[52]

These dialogues, in my reading, mark a third mutation in the relation between Habermas' formal answer to the question of modernity and religion. As I will show in this section, Habermas abandons the idea of a difficult and unclear coexistence to embrace a new project: defining the limits of a possible collaboration guided by principles of translatability.[53]

The theological language, in Habermas' understanding, could be the bridge that makes collaboration possible. Contextually to the study of the contemporary theology, Habermas explores the possibility for both secular and religious citizens to find a common language characterized by the combined effort to translate religious and secular reasons in a language understandable by all the speakers.

Once it is recognized that within a post-metaphysical philosophy "the very syndrome of validity" of the privileged metaphysical and religious views simply dissolved without any space for the reconstruction of a global horizon of meaning, not even a negative one, then "what remains for philosophy is an illuminating furtherance of lifeworld process of achieving self-understanding"; yet this very process of self-understanding "must operate under condition of rationality that it has not chosen", "it is limited to grasping universal structures of lifeworld in general".[54]

---

[52] Habermas' essays *Communicative Freedom and Negative Theology* and *Israel or Athens: Where does Anamnestic Reason Belong?* are probably the highest moments of this dialogue. They are both published in Habermas, *Religion and Rationality*, ed. E. Mendieta, Cambridge, Polity, 2002.

[53] Habermas passes from the prevalent use of the term 'appropriation' to the one of 'translation' without explaining the shift, or better the enhancement, in his terminology. Yet, it is possible to see that such a passage is congenial to his more recent recognition that a) discourse ethics cannot critically appropriate contents that do not lend themselves to the kind of transformation that makes them critically appropriable, and b) the needed transformation is possible only on the ground of a collaboration between those who experience those contents and those who intend to use them.

[54] Ibid., pp. 17-18.

A direct consequence of the interdiction of meaning from the sphere of critical thought is the necessity to look for it somewhere else. Religion, or better the formalized Jewish-Christian tradition that is freed from the burden to prove its rationality appears like the most convenient—and extraordinarily similar to what Habermas had hoped to obtain with discourse ethics itself—source to use in order to preserve a minimal unity of the social fabric which, in that way, would be able to provide the necessary motivation to interact in line with communicative reason. Those fragments of tradition, in fact, do not require to be artificially added to the cultural substratum of the so called Western societies, or created *ex nihilo*; they are already part of it. All that is needed, in order to foster their unifying and motivating role, is to bring them to a self-conscious level. This, according to Habermas, is precisely the "illuminating" work of philosophy.

In an essay from 1991 against Horkheimer's thesis about the impossibility of maintaining a concept of unconditionality without God, Habermas clearly states that philosophy and all the post-metaphysical thought differ from religion because the former is able to save the unconditioned without any proper idea of the Absolute.[55] This unconditioned owns its status to a form of transcendence—the normative structure of language itself—to which "whoever employs language with a view to reaching understanding lays himself open", so to speak, "from within". According to Habermas, such transcendence "from within", which is understood as in opposition to the metaphysical/religious one "from without", suffices on its own to overcome historical and cultural limits without forcing any of the competent users.[56]

[55] Habermas, "To Seek to Salvage an Unconditional Meaning Without God Is a Futile Undertaking: Reflections on a Remark of Max Horkheimer", in *Religion and Rationality,* ed. E. Mendieta, Cambridge, Polity, 2002.
[56] Ibid., p. 108.

Despite the stated autonomy of the transcendence "from within", the limits which divide it from the one "from without" appear more porous than they were in Habermas' precedent works. When, in the pertinent paper, he moves his attention from the formal aspects of his post-metaphysical uncoditionality to the "unconditional meaning" of religious contents, time and again, he associates it with consolation, and therefore emphasizes the necessity to exclude that form of unconditionality from the sphere of philosophy *tout court*. Such exclusion, though, is now a conditional one:

> it is altogether a different matter to provide a motivating response to the question of why we should follow our moral insights or why we should be moral at all. In *this* respect, it may perhaps be said that to seek to salvage unconditional meaning without God is a futile undertaking, for it belongs to the peculiar dignity of philosophy to maintain adamantly that no validity claim can have cognitive import unless it is vindicated before the tribunal of justificatory discourse.[57]

What Habermas is doing here is not as smooth as he suggests. On the one hand, he seems to underestimate the radicality of the question about why be moral, and the relation this question bears on the process of justifying a moral point of view (despite the fact that these are undeniably two different issues); on the other, he seems to open up to an integration of rationality with—also—the very consolatory elements provided by religious languages.

The now possible vindication of those "indispensable" contents was, in fact, considered impossible until at least 1988.[58] After that period, the religious sphere is considered, with an increasing attention, as possible preferential interlocutor compared with other competitive secular traditions, for it provides means to deal with the complexity

[57] Ibid., p. 108.
[58] See previous sections about Sublation, and Coexistence.

of modernity, namely, the kind of identities that discourse ethics needs, that is to say citizens who are able to both oppose the decline of the Jewish-Christian ideas of equality and justice, and allegedly willing to display the degree of tolerance necessary for Diskurs.

As I mentioned at the end of the previous section, Habermas, during the years that led to the third transformation of his understanding of the relation between philosophy and religion, was preparing a classical Hegelian move. The religious tradition Habermas refers to, starting from the 90s, appears as having its own degree of rationality that can be brought at a reflexive level. Such a reflective stage, not only can, but has been reached through the dogmatization of the contents of the experiences of faith.

Theologians, and all those who have reflectively engaged with religion, have been able, according to Habermas, to develop languages that are “hermeneutically competent”, and whose relation with subjects is not objectifying.

One of the clearest and most advanced forms of such an elaboration of religious contents is provided by the German theologian Mets. His effort to comprehend and make accessible religious experiences otherwise resting at an unreflective, and therefore non argumentative level, is considered by Habermas as an enormous source for the kind of disposition necessary for discourse ethics.

However, this means that the point at which religion can make contact with reason is not through the religious experience as such, but just through the mediated self-comprehension of the contents of faith. The experience of faith, in fact, if not elaborated at an argumentative level, ends up, according to Habermas, facing a fate similar to the one of political theology. That is to say, the absence of a translation from specific languages to one

universally comprehensible makes the collaboration between discourse ethics and religion impossible, for those languages are "no longer directly measured by truth claims".[59]

Philosophy, Habermas writes, "cannot appropriate what is talked about in religious discourse *as* religious experience". When, regardless of this impossibility, the philosophical discourse pushes itself in this extraneous field, it turns itself into literature and religion in "mere citation".[60]

Despite this danger, Habermas seeks to draw precisely on religious languages. Thanks to the dialogue with Metz, and Theunissen, Habermas shows the mature signs of this third mutation.

The idea shaping those dialogues is to find a common ground for a critical exchange that avoids the dangers mentioned few lines before. In order to make this possible, Habermas outlines a series of methodological considerations concerning the reciprocal historical influences; influences that, nonetheless, mark something that does not cease to be an alterity. These Habermasian preliminary historical and methodological considerations appear like a sort of prophylaxis—a prudent restatement of the reciprocal limits that is made in order to avoid unwanted contagion.

Even if prudent, the essay that Habermas commits to the analysis of Metz's work is the prelude to a dialogue between reason and religion. This is at least the way Habermas conceives it, even if he pursues this dialogue in the very narrow space that opens up between receptivity and methodological scepticism.

Productively challenged by his interlocutors, Habermas tries to delineate an argumentative justification for a *docta spes* which extrapolates from a reasonable pessimism reasons to believe that "everything with time will be different", without falling

[59] Habermas, *Religion and Rationality,* ed. E. Mendieta, Cambridge, Polity, 2002, p. 74-75.
[60] Ibid.

into the unjustifiable belief that "the time itself will be different".[61] For Habermas, there is no place for this eschatological dimension within secular rationality. The image of a completely emancipated humanity is confined to an inaccessible otherness made up either of outdated metaphysical dreams, or of blind *spes fidei*.

The necessity for a non defeatist reason, in addition to Metz argumentative effort, pushes Habermas to the core of religious issues without being able to critically appropriate them—this for reasons I will focus on later on in this section.

Metz, with what he understands as a de-Hellenized conception of time, stresses the fundamental role of memory in every discussion about justice. The concept of progress, according to Metz, cannot provide any liberating force. It, by forgetting the negative, removes what has been in order to proceed towards a future qualitatively constant that cannot but reproduce injustice exactly because it does not take on past injustices. The core of a liberating reason, for him, is anamnestic. It is anamnesis that can bring justice to the "wreckage upon wreckage" that, for Benjamin, history has piled up. This kind of reason—as a repository of a mystical force, through the exercise of a memory that is able to bring to consciousness all past human injustices—is, according to Metz, the only possible way to retroactively reconcile, across present, past and future.

Habermas is not indifferent to Metz's analysis of the radical demand of justice, but the closer his reflection gets to the question of meaning, the more evident appears the fragility of the "transcendence from within", and therefore more pressing becomes the necessity for him to protect what seems the only rational possibility for the unconditional—discourse ethics.

[61] Ibid., p.133.

It is this necessity that leads Habermas to avoid the direct philosophical engagement with concepts of hope, humanity, and emancipation. At the same time, however, the semantic power of those concepts—even if they are not fully rationally appropriable—appears as necessary for the full functionality of a profane justice that can count only on the public opinion of citizens.

Discourse ethics, now divested of the original project to completely replace the old objectivity, seeks an alliance which, however, it cannot fully embrace. Habermas seems to expect, given the impossibility for discourse ethics on its own to appropriate the contents it needs, from religion a spontaneous step towards argumentative language. But the fact that Habermas sees the necessity of a convergence between cognitive sources and the normative intersubjectivity that would make discourse ethics feasible does not make such a convergence likely. For, neither philosophy, nor any rational procedure, can impose such convergence.

Habermas once again finds himself in a position for which he has the perfect tool, a tool reconstructed from the formal-pragmatic functioning language of an actual society of speakers, yet what he would actually need in order to make discourse ethics work is a counterfactual society—the one willing to cooperate.

During the last ten years, the validity of religious contents, which in the 90s was just mentioned, has become a relevant part of his entire work, forcing Habermas to rearrange, or at least to relativize the equilibriums of his evolutionary conception of reason. The new change does not modify the ambiguity of Habermas' interest in religion which swings between its social function, and reflections *around* its meaningful core.

One of the key texts of this period is the acceptance speech as recipient of the Peace Prize of Frankfurt in the 2001, called *Faith and Knowledge* in which Habermas brings

together the principal themes of his work on religion, providing in this way a useful and coherent perspective.[62]

Apart from the initial mentioning of the instrumentalized use of religion, Habermas focuses on the "cooperative task" of the, now, post-secular society. The project of secularization is taken as overcome, the new historical context (recently Habermas has started to talk about a post-secular time) "is prepared for the continued existence of religious communities in a continually secularizing environment", as if religion were a sort of inextinguishable source. And what is even more surprising in the light of Habermas' previous views: "the boundary between secular and religious reasons is fluid".[63]

What is prefigured in the speech is a possible fruitful exchange; an exchange that is first analyzed against the backdrop of the question of secularization. Habermas holds that the two main opposite interpretations of secularity are equally wrong, both the "outstanding", and the "expropriation model" are unable to see that the transfer of authority from religious institutions to the secular state in itself is neither a cause of infinite progress, nor a disastrous misappropriation that will lead moral society towards into an abyss.

The initial passage of authority is, instead, part of a transformation that has effected, and still does, profane and sacred elements of the post-secular society, leaving space for the emergence of a "democratically enlightened commonsense", which, as sort of "third party" plays its "civilizing role". The porosity between the two kinds of rationality brings with it the necessity for a sort of reciprocal responsibility—a further transformation.

[62] A perspective even more inclusive about Habermas' work on religion has been recently provided by a collection of essays edited by Mendiated and enriched by Habermas' own answers to each of the works included.

[63] See: http://homepage.newschool.edu/~Wilhelm/Fall2005/Faith%20and%20Knowledge.pdf

In order to save secular reason from an "evacuating deflation", Habermas suggests a "critical adaptive transformation of the religious content", a transformation that is not only the one that already has historically happened (the conceptual Jewish-Christian heritage of the entire western thought), but a constant one whose necessity imposes itself when moderns feel "to owe each other more".

The evocative tone, the use of expressions like "the unsentimental desire to have the suffering inflicted on others undone" provide a sort of ethical break-in, for if there is a form of desire which is not an affect, then, it must be a fundamental ethical orientation.

This necessity of a semantic alterity that discourse ethics has, is presented as a sort of inexplicit ethical dimension of the speech itself. It, once more, highlights that meaning is something that, despite a life-long attempt do to so, does not lend itself to be reduced to intersubjective webs of free discoursivity. It is as if Habermas had discovered during the last years that people do not only speak, but they also die, and that it is just the most extreme of the speechless experiences that impose themselves with an open demand for meaning.[64]

Habermas' unprecedented turn to the ethical sphere does not mean that he reshapes discourse ethics in order to make room for any question about meaning itself, rather, he tries to open a preferential channel for those kinds of traditions which, according to him—and before any validation process—have proved to better support its functioning.

The fragments extrapolated from the Genesis that Habermas quotes are just the rhetorical support for a Diskurs that on its own will be unable to work. It is Habermas himself who recognizes the insurmountable limits of a procedural rationality and seeks a collaboration which does not move away from universality. What Habermas is suggesting is a division of labor, the division between an objectivity which takes on itself the

[64] About this point see Habermas essay "An Awareness of What is Missing" published in the homonymous book (Cambridge, Polity, 2010).

responsibility for the experimental creation of a potentially universal self-narration that is able to cope with the changes imposed by instrumental reason, and communicative rationality which supervises from within the validity claims of those narrations.

It is in defense of the normative sources that would disappear into a completely reified lifeword that Habermas establishes difficult dialogues even with, at the time, cardinal Ratzinger. During the 2004, on the occasion of a rather bizarre 'dialogue', Habermas does not let the opportunity pass to recalibrate the position of discourse ethics, and consequentially of the deliberative democracy in regard to its pre-political sources. [65]

The cue is the well known thesis of Böckenförde formulated during the 60s: the liberal state has its roots in an ethical substance (religious tradition included) which is pre-legal and pre-political, therefore "the fact of plurality" can be integrated cognitively, but not normatively.[66] Habermas, after a short look to other classical Böckenförde-sympathetic accounts of the Nation-State such as Jellinek's and Schmitt's, clearly states that sovereignty, in a constitutional state, cannot draw on a pre-legal substance.[67] It is equally true, though, that neither rationality, nor "political virtue" can be legally imposed.

Habermas tries to solve what is, according to him, just an apparent circle. Political virtues, Habermas explains, are in a certain sense pre-political, in addition they are also "a matter of socialization", though they do not hold a precise content; rather they are a modality, "a mode of thinking".[68] Habermas' strategy is "to give a defuse interpretation of Böckenförde theorem" by making the ethical substance reflective, and therefore translatable into concepts that are accessible also to those who do not share that original

[65] Habermas, *On the Relations Between the Secular Liberal State and Religion*, in Political Theologies, edited by Hent de Vries and Lawrence E. Sullivan, Fordham University Press, New York, 2006. From now on RBSR.
[66] E. W. Bockenforde, ''The Rise of the State as a Process of Secularisation,'' in his *State, Society and Liberty: Studies in Political Theory and Constitutional Law*, trans. J. A. Underwood (New York: Berg, 1991).
[67] RBSR, p. 251.
[68] Ibid., p. 253.

cognitive stance.[69] At that point, concepts and motivations belonging just to a part of civil society are ready to be validated through public argumentation. This means that the collaboration between discourse ethics and a precise ethical stance, religion *in primis,* is never a direct one, it must be mediated by translation. What is essential for the possibility of such a translation is a favorable disposition, what Habermas defines as a "learning process", which rests on an almost obvious "epistemic asymmetry".

While religious beliefs, as every other object of faith *qua* faith, do not require any form of validation, public argumentation is open to its own fallibility, and in need of constant validation. This difference, according to Habermas, must be pointed out before an actual exchange can happen. That means, a change in the perspective of believers concerning their own convincement is required.

Religious people are not the only citizens from whom a change of perspective is expected in a post-secular society; non believers have to change their perspective too. This mutation into a reciprocal interaction is underpinned, in Habermas' eyes, by his socio-political analysis of post-secular societies. Still in accordance with the diagnosis formulated thirty years before in *Legitimation Crisis* Habermas restates:

> that the balance that has emerged in modernity between the three major media of societal integration [power, money and communication] is now threatened because markets and administrative power drive societal solidarity […] out of ever more areas of life. It is therefore also in the constitutional state's own interest to treat with care all cultural sources upon which the consciousness of norms and the solidarity of citizens draw.[70]

[69] Ibid., p. 258.
[70] RBSR, p. 258.

This instrumental view of semantic sources is apparently completely reasonable, and however comprehensible within the economy of Habermas' procedural approach, it shows, nonetheless, all its ambiguity, and the difficulties of discourse ethics. It is condemned, on the one hand, to be crossed by meaning without being able to engage with anything but its valid form, and on the other hand, it needs to pre-suppose meaning endowed with specific qualitative characteristics in order to work. The religious pre-comprehension with its non conscious condition (Lebenswelt) belongs to the community of free beings; even if in its reflective form it is necessarily linguistically mediated (Weltanschauung), it is also a presupposition of a non distorted use of the communicative reason (Diskurs).

This makes pre-comprehension an essential complementary feature of discourse ethics. This would not be an insurmountable obstacle if this condition of complementarity was 1) equally open to any (i.e. also non-Judeo-Christian) pre-comprehension, and 2) not unbalanced in favor of the irrational (in so far it is unspeakable) pre-comprehension, as it seems to be at this stage of Habermas' analysis of religion. The possibility to question the very substantial reasonabless of the pre-political "mode of thinking", is, in fact, banned from Habermas' post-transcendental critique. This impossibility translates in Habermas being forced to identify a 'good' semantic source before even checking the validity of its claims. With the aggravating circumstance that Habermas uses the fact that religion no longer has a claim to epistemic validity as a reason to avoid facing this structural problem of discourse ethics.

Paradoxically, this means that Habermas' answer to the question of modernity is, despites his claims to the opposite, relying on a sort of faith in the non-validated objectivity (religious beliefs in this case, but not only).

The reliance of discourse ethics on a disposition to enact and protect it, implies a reliance on not validated meaning nullifies Habermas' project to provide a fallible unconditional. His neutral and formal standpoints, supposing it could work – or better, supposing it finds in place the conditions (the dispositions) it presupposes to work – gives up the possibility of critically assessing those features on which its proper functioning depends.

This difficulty brings Habermas back to the starting point of modernity—Kant. In the same year of the dialogue on the relation between the liberal state and religion, that is in 2004, he has to reconsider *The Boundary Between Faith and Knowledge*[71]. This reconsideration, which is a kind of retracing again the steps that from Kant, through Hegel, Marx and Kierkegaard, have led him to his own critical incorporation of a religious disposition— a sort of "rational faith"—, is a way to stress once more the limits of a non-religious philosophy that aims to critically "rescue the contents of faith".[72]

This critical appropriation is seen as necessary to support a non defeatist practical reason that, while not allowing itself any comforting idea of final happiness, still interprets the "lawful necessitation of action", which is separated from the specificity of the ends, as open to the possibility of "the final end", that is to say, the realm of ends.[73]

Since the condition of possibility of the not-yet existing ideal community is still something that eludes both the limits of reason and those of duty: unselfishness is no way secured.[74] How to counter balance the phenomenon of selfishness radicalized by self-alienating autonomous subsystems of modern time is the question which works as engine for Habermas' interest in the role of religion within the limits of its linguistic translatability.

[71] BNR.
[72] BNR, p. 221
[73] Ibid., pp. 219-220.
[74] Ibid., p. 220.

The research into the specificity of the religious form of self-comprehension remains something radically extraneous, as it is extraneous to morality any ethical question about the very motivation of a moral disposition, though Habermas, in the case of religion, reckons as possible to neutralize its ethical content, by translating it. The process of translation would make semantic content highly reflective, keeping at the same time unaltered its power to shape and maintain what, according to Habermas, is merely a non egoist disposition. In other words, Habermas, rather than just engaging in an actual dialogue with Kant, his practical philosophy, and his critical study of religion, employs a Kant that is partially Hegelianized to transform an analysis of the possible substantial condition of validity into the tasks of taming a cognitive disposition.

The evolution of Habermas' positions in relation to religion not only did not find a way to solve the "curious dependence" that discourse ethics has, but it also produced, in each of its formulations, new problems. Once the process of sublation of religious contents and of the role of God showed itself incomplete, Habermas was already dealing with the motivational problems of his most refined answer to the question of modernity. This condition pushed him, on the one hand, to recognize the coexistence between religious contents and fully secularized reason. On the other hand, the expressive and motivational force that he recognized in those contents forced him to consider them useful but non appropriable by discourse ethics. This made clear that discourse ethics was exactly in the uncomfortable position in which Horkheimer's critique once was, that is to say, being a necessity that is unable to turn itself into a feasible possibility. At that point, Habermas thought that that problem could be overcome thanks to the mediation of a third not too dissimilar from either religion or philosophy: theology. The division in three different elements (religious experience, reflexive religious contents, and the formal process of post-

metaphysical thought) disclosed, according to Habermas, the possibility of a fruitful collaboration. A change in the cognitive disposition of both believers and non believers was required. This change would have made possible the translation of the respective reasons in something commonly sharable and therefore open to public discussion (discourse ethics).

This, seen from the believer's perspective, seems less like a collaboration and more like an invitation to give up faith in exchange of a secularized understanding of certain beliefs. In addition, those beliefs, once translated in a reflective form, will be appropriated (because in this form they can be appropriated) by argumentative cognitive forms. First an abdication and then the surrendering of all the goods (the useful rational contents) to what pushed for that renunciation in the first place (discourse ethics). This looks like the kind of collaboration from which one does not recover.

Seen from the perspective of radical critique, this strategy also amounts to giving up. Not only does discourse ethics implicitly acknowledge its dependence on specific semantic contents that it cannot assess or validate (because the dependence makes this impossible), it, in this way, depends even on the believers' will to give up on status of reveled truth of some specific contents of their faith. That is to say, the collaboration between religion and discourse ethics, which is so precious to discourse ethics, is unbalanced in favor of religion. And it is unbalance in a manner that is particularly disadvantageous for discourse ethics because, at this point, it must count on the will of religious people to give up, or at least, share their faith.

What is more, even if believers and non-believers would find a manner to collaborate in the way Habermas auspicates, as Cooke points out in several of her works,

such a collaboration would still bear methodological requirements that are at odds with discourse ethics' proceduralism.[75]

As Habermas himself made clear over the years, his attention to religion does not exhaust itself in the study of the historical learning processes that brought about the historical conditions of possibility of discourse ethics, as, for instance, the affirmation of concepts such as universal justice and equal dignity of every and each human being. The knowledge about how we moderns got here, for Habermas, is not the main issue when it comes to his engagement with religion.

What Habermas is interested in is motivation, and inspiration—in one expression—alive meaning. And this is particularly clear, as I pointed out in the section entitled 'Coexistence', when he focuses on the experiential element of faith—an element that, at that stage, is considered both a reservoir of "indispensable" meaning, and inappropriable precisely because of its experiential character; a problem (the appropriation of a reservoir of meaning) that does not emerge when it comes to the already reflexive pieces of knowledge, as the hope for a universal emancipation, which are already part of the history of modern thought.

When Habermas, subsequently, tries to overcome this obstacle by both calling on believers' (and non-believers') collaboration, and introducing the idea of a possible translatability of religious contents in a language accessible to everybody, he, again, speaks of a learning process. Yet, Habermas is no longer talking of an historical learning process. He is referring to a different process capable to mediate, and consequently make religious experience reflexive.

[75] See Maeve Cooke, 'The Limits of Learning: Habermas' Social Theory and Religion', in *European Journal of Philosophy*, 2014; and by the same author 'Translating Truth', in *Philosophy and Social Criticism*, 2011, pp. 479-492.

It is at this point that Cooke notices that if we are all going to learn from those religious experiences then there must be something that we recognize as true there. Yet the kind of truth that we recognize in those experiences, when and if we recognize it, is not a procedural one. She defines it as "truth as disclosure".[76] This is a definition to which she arrives developing her research about what translation means.

Cooke points out that an adequate account of what translation means "requires us to keep sight of the question of truth".[77] However, when it comes to religious contents and their exemplarity, the possibility to be hooked by them is opened up by a non-argumentative element. That is to say, in order to be receptive to what is extraneous to our way of thinking and seeing, there must be something "that resonates with the subjectivity of the addressees" something that "opens their eyes to new ways of seeing the word and themselves". This experiential element that Cooke defines as "truth as disclosure" is what makes translation capable to re-present the truth of the original.[78] Now, even if she sees this kind of truth—which is based on personal experience and context—only as "co-determinants of rational acceptability", and therefore as something that can be integrated in post-metaphysical validity claims, it remains the fact that in order for something to be translatable, the truth as disclosure must be directly engaged.[79] That is to say, Habermas should develop analyses capable, and willing, to take into account experiential and context related—and not only language-immanent—forms of validity.

Instead, Habermas, by entrusting the practical field and its legitimacy exclusively to linguistic communication has impoverished the notion of validity to the point of not having

[76] M. Cooke, 'The Limits of Learning: Habermas' Social Theory and Religion', in *European Journal of Philosophy*, 2014, p. 8.
[77] M. Cooke, 'Translating Truth'. *Philosophy and Social Criticism*, 2011, p. 481.
[78] M. Cooke, 'The Limits of Learning: Habermas' Social Theory and Religion', in *European Journal of Philosophy*, 2014, p. 8.
[79] See in particular footnote 10 of 'The Limits of Learning.'

space for the practice of translation that would be required for his purposes, as he has no instruments to deal with the ethical and aesthetic dimensions.

The repeated attempts of fixing the workings of discourse ethics by integrating elements that discourse ethics needs and relies on, without being able to actually integrate them, show that Habermas has indirectly given up on his original project to revitalize the spirit of modernity by way of a pragmatics that provides an unconditional standpoint in virtue of which reason can scrutinize its own ends. Furthermore, this implicit renouncing takes a sort of theatrical turn. The involuntary renouncing of normative justification as his answer to the question of modernity brings Habermas closer than he would like to those critical theorists who renounce since the beginning—and voluntarily so—the quest to establish normative foundations–I am thinking of Foucault, for instance. However, at this point, I want to shift my focus on another battle ground in which we can see Habermas implicit renouncing radical critique: the self-assertion of discourse ethics in the face of the self-objectifying power of bioengineering.

Gattungsethik—an ethics of the species

As pointed out at the beginning of this chapter, Habermas' intellectual project of providing an answer to the question of modernity places itself within a clear path. In the wake of the first generation of Frankfurt School, he tries to provide a non religious and non metaphysical foothold for a radically critical practice. This ambitious project is formulated as answer to a manifold crisis of the 'objective' that has accompanied the emergence of a radical form of self-reflexive reason which is specifically modern. This project has substantiated itself primarily in the shaping of a formal pragmatics that, drawing on the

inescapable normative structure of linguistic reason, is allegedly able to justify itself independently from any conception of the *good*. In other words, Habermas, in order to provide a universal tool for democratically, and yet rationally, assessing the aims reason gives to itself, has restricted his intellectual battle field, so to speak, to the sphere of the *right*—a theory of justice.

So far, and in a way that is still open to refinement, I have shown that this theory of justice, as soon as it is put into practice, has faced structural difficulties in dealing with meaning, and in particular with religious contents.

In this section, I focus on what happens to discourse ethics and the whole Habermasian critical project once he, pushed by recent evolutions of biotechnical knowledge, shifts his perspective, concentrating on how to avoid radical evil under these new conditions.[80] A radical evil is, according to Habermas, what does not amount to an infringement of shared moral norms (wrong), but nonetheless threatens to undermine morality itself.

This is for me of particular interest because this change of perspective brings out problematic presuppositions of discourse ethics that otherwise go unnoticed, as I point out in my assessment of Habermas' critical project. I will demonstrate that the ethics of the species in which discourse ethics is grounded is to be regarded as highly problematic given Habermas' own requirements with regards to what counts as an acceptable answer to the question of modernity.

[80] At p. 110 of *The Future of Human Nature* (from now on FHN), Habermas clarifies that we still lack an adequate comprehension of the semantic difference between what is morally wrong and what is profoundly evil, yet as we will see, he seems quite confident in this work about what counts as evil.

The work in which this change of perspective can be best observed is *The Future of Human Nature*. A text in which Habermas tries to articulate his gut reaction against the idea of engineering the genetic code of future human beings.[81]

A fair amount of second literature has started to be available now about this work, yet, to my knowledge, they focus almost entirely on i) the fact that Habermas lacks arguments to support the statement that autonomy and equal moral standing of future humans could be seriously and irreversible hindered by positive prenatal genetic engineering, and ii) the presence of a fundamental problem concerning what would count as a change in the human nature.[82]

I want to argue that even if critics are right that Habermas' arguments fail to support his view that autonomy and equal moral standing would be undermined by enhancing future people, it is still the case that we can learn something from this work about Habermas' failure to give a convincing answer to the question of modernity. This is because we have to distinguish between three issues that Habermas addresses: autonomy, species ethics (*Gattungsethik)*, and the instrumentalization of inner nature.

Instead of focusing on the first of these three issues, I will be mainly concerned with the second, because I want to demonstrate that it plays an indispensable (because complementary) role for his theory of justice. Consequently, it is precisely the concept of

[81] FHN, p. 22. "this essay is an *attempt*, seeking to attain more transparence for a rather mixed-up set of intuitions. I am personally far from believing that I have succeeded…"

[82] About the first point see Harris (2007) or Agar (2004). They both base their 'pro' argument on the conviction that maintaining a sharp distinction between technology and nature is already impossible. If the distinction is already blurred then a genetic intervention will not insert any change in that blurred relation, therefore the self-understanding of human species is not in danger. Regarding the second point, Buchanan (2011) points out that the idea of some sort of human biological essence endangered by medicine is ungrounded and that the very concept of human nature is so confused to be of no use. On a similar argumentative line, Lewens (2012) who adds that the concept of human nature is not only useless but dangerous. Christiansen (2009) is an interesting exception which focuses on how Habermas allegedly fails to show the relation between his existential analysis and the sociological and psychological impacts of genetic enhancement.

Gattungsethik that provides me with a way to investigate how discourse ethics relates to, or presupposes a minimum ethics.

Glimpses concerning where the Habarmasian horizon was turning were perhaps already visible in the increasing attention he was, and still is, devoting to the theme of religion and to authors who have given within their reflection wide space to questions that go beyond the concept of justice understood as giving equal consideration to interests.[83] Yet, the way in which Habermas has turned his attention to ethical questions has happened in a rather unexpected manner. Habermas allows himself to speak about ethics outside the limits of his theory of justice when, as eminent and globally recognized intellectual, he decided to exercise his public role to express his worries about eugenic practices.

On that occasion the references to religion are minimal, but the weight of the thought of a religious thinker, Kierkegaard, is vital for the argumentative economy of the whole work. It is as if the appearing of a radical danger made possible by recent biomedical developments had made Habermas realize some specific limits, or conditions of possibilities, of discourse ethics.

It is not the case that Habermas unexpectedly develops a philosophical interest in what counts as a good life. That, insomuch as it is a strong ethical concern, does not enter in the new Habermasian horizon. Yet he turns to a form of species ethics *(Gattungsethik)* because he realizes that such an ethics makes morality possible (and by implication discourse ethics). Even if "philosophy […] investigates only the formal properties of the processes of self-understanding", the very possibility both of being a self and to understand this self as free is understood by Habermas as enabled by the fact that each self can only become a self against the backdrop of a substratum of unconditionality.

[83] What I have in mind is, for instance, the fascination Habermas has for concepts such as anamnestic justice.

Genetic engineering gives to that substratum the mark of human will. This compromises exactly that unconditionality by introducing a dangerous difference within the community of human beings. The danger, according to Habermas, rests on the fact that this difference produced at the ethico-biological level could be echoed at the level of the moral community and destroy it.[84] For a moral community, in Habermas' sense of it, is either composed by beings that share the status of equals, or it does not exist at all.

It is thus not a concept of the good as such that sustains the development of Habermas' reasoning against positive eugenic. Instead, it is the necessity to establish a weaker ethical point of view that seeks to protect the conditions of possibility of morality. Habermas therefore becomes aware of those conditions and tries to identify what would undermine them.

The modification of the species, and therefore of the form of unconditionality it stands for, is the radical evil to avoid. It is for Habermas necessary to leave the species untouched so that (i) the horizon of possibility of each human being remains equally open, and (ii) the status of free and equal humans, on which the right is grounded, remains safeguarded.

The way in which Habermas articulates his negative reaction to positive biological engineering and the limits of *Gattungsethik* are the subject matters of the remainder of this chapter.

[84] I speak here of an ethico-biological level because, as we will see later, *Gattungsethik* is a mix between sharing the genetic code that characterize the species and recognizing oneself and the others as beings who share that undisposable common ground.

From Kierkegaard to Gattungsethik

The analysis of what is fundamentally evil about genetic engineering is preceded by a philosophical appropriation. Kierkegaard's ethical theory of subjectivity is employed to voice Habermas' own understanding of how, and under what conditions, a subject is able to choose herself. Drawing on both *Either/Or* and *Sickness unto Death*, Habermas explains that the passage from an unsustainable aesthetic mode of being to an ethical one that is the one characterized by the possibility to be oneself, necessitates coming to terms both with the world and with oneself.

Choosing oneself is neither a mere exercise of will nor a simple pain-free process. The person that one decides to be is not open to the realm of logical infinite possibilities; rather, it is only possible to carve out this person against a backdrop that is itself not chosen. Yet the very choosing, under these conditions, is the source of the responsibility that one bears towards oneself, and towards the project of life one is.

In order to be able to embrace herself, the ethical being needs to accept and recognize the fact that what enables her as person is something other from her—something radically Other. In Kierkegaard's case that transcendental power is God.

According to Habermas, this process of individualization that Kierkegaard describes is "sufficiently formal" to remain relevant for a post-metaphysical thinking. It is particularly suitable for being used in order to explain the role Habermas gives to language in his own understanding of the enabling conditions of subjectivity.

The transcending, yet enabling, power that is beyond each individual is language itself for Habermas. The possibility of understanding oneself as an autonomous subject,

and, at the same time, the chance of coming to an understanding with other subjects are given by this intersubjective other which is virtually inescapable.

As rational beings, according to Habermas, we all depend on language, and language is what establishes us as rational beings. Even if language is a human product, so to speak, there is no human being that can decide the fate of language as whole. In this respect, language is the undisposable condition that enables us as rational being.

However, if I understand Habermas properly here, what he points out in *The Future of Human Nature* is that language is not a sufficient condition for reaching the status of *free* and equal moral being. There is, in fact, another precondition that he has not addressed in his previous works—the species.

Without actually articulating the relation between the unconditionality that language is and the one embodied by the naturally, that is unintentionally, developing of the genetic code, Habermas simply implies that both forms of unconditionality must be in place in order for morality to exist.

The status of free and rational derives from the fact that every single subject comes to understand herself against conditions of unconditionality, that is, conditions that do not bear the mark of any will, or of any particular conception of good, at least not the kind that cannot be critically rejected. After all, even if the process of individualization via the medium of language is embedded within a specific lifeworld, the latter is in turn linguistically shaped, therefore, according to Habermas, it is open to "revisionary learning processes". Instead, the possibility of disposing over the inner nature of human beings would irreversibly mark the *status* of future humans.

The manipulation of the genetic code would undermine the possibility of being-able-to-be-oneself in an existential manner that could then be mirrored at both the psychological and social level.

This possibility of not identifying and not being identified with the rest of the human consortium, in Habermas understanding, does not necessary imply the inability to make rational choices. What is undermined is a rather wide or simply different conception of autonomy.

Autonomy

Once Habermas has appropriated Kierkegaard's formal structure of the process of individualization, stressing the essential role of unconditionality, he ventures in an attempt to explain why and what would be specifically undermined if, as a community, we would start to enhance embryos. This explanation is repeated in the text several times and from slightly different points of view, but it can be reduced to three essential and interconnected consequences deriving from the voluntary modification of the human genetic code. This modification, according to Habermas, would mean: i) the introduction of an insurmountable disparity between humans that will push people genetically modified to consider themselves as not autonomous but as a (non)self designed according to some other person's conception of a desirable being; ii) the introduced diversity could, at one stage, undermine the sense of a shared humanity, and iii) the objectification that would derive from it would be the most perfect embodiment of instrumental rationality. The same kind of rationality which has first destroyed reason as stable moral agency (at least according to Horkheimer) and that now hinders its renewal.

The first point, the one concerning autonomy which most directly links up with the issue of unconditionality is probably the theme that has been more discussed by Habermas' critics and there is actually wide space for various kind of objections there. This is probably also the case because the term autonomy is employed in a surprisingly (since we are talking about Habermas here) ambivalent way.

When Habermas states that if parents start to modify the genetic code of their offspring, the people that will derive from that intervention will not share with the rest of community "the spontaneous self-perception of being the undivided author of [their] own life", one has the justified impression that what Habermas is talking about is accountability. This impression is further justified by the explanation of what would happen in cases in which a person is no longer able to recognize herself as fully responsible for her entire life. Habermas writes that the consequence would be that enhanced people will find themselves with interpersonal relationships "no longer consistent with the egalitarian premises of morality and law". The egalitarian premises, in fact, would be already gone as soon as there would be people who have chosen themselves (in a Kierkegaardian sense) and people who have been irreversibly influenced qua genetic programming by a conception of the good life which is not of their choosing.

Yet if one thinks about the Kierkegaardian preface, accountability seems not to be the issue at stake, or at least not the main one. If it were, authors like Harris would easily hit the target when they point out that the past will of parents has a huge influence on their children even when it does not take the shape of genetic engineering, yet this influence in itself does not undermine the autonomy of children. Children have not given their consent to the kind of diet the mother has followed; they have not picked the schools they have attended, and so on. The autonomy of these people has emerged from an

inextricable mix of alien wills and in this way they are equal to all the other human beings.[85] For this reason Habermas' division between the chosen and the made then is a very weak argument since it is not possible to prove (or at least Habermas does not prove) a causal influence of the genetic intervention for the kind of person one will be.

But in my reading, Habermas is actually grappling with a slightly different kind of issue. Without negating the evident presence of a discussion about what makes one the author of one's own life, it seems to me that the question that Habermas is trying to face is what makes one the author of one's own *human life*. At least this is the only way in which I can makes sense of the attention Habermas places on the importance of not disposing over the biological features of a future person.

What is at stake is the status of human being, not rationality, not the ability to make decisions. There is no point at which Habermas states or implies that enhanced people would grow exactly according to the idea of the good that guided their enhancement, as a sort of automata. Even when reflecting on autonomy, Habermas seems to be dealing with identity; not the formal process of acting according to a certain normative idea of the self, but the pedigree of that idea. Habermas repeats over and over again the verb "to recognize": recognize oneself, recognize the others, and recognize the community as such. Failing to do so does not undermine autonomy, instead it undermines human autonomy.

It may seem an absurd way to differentiate, and it may well be, but Habermas seems to suggests that so far the only kind of autonomy we have known is the human one; genetic engineering opens new scenarios in which one can be autonomous but not in the way we have known it. Therefore the concern for Habermas is given by the fact that the moral community can not survive to that difference.

[85] Harris, J., *Enhancing Evolution*, Princeton University Press, New Jersey, 2007.

If I am right, "the egalitarian premises" Habermas is trying to keep formal by focusing on the *process* of self-understanding and self-choice are not formal at all, and it is for this very reason that he needs to introduce the concept of *Gattungsethik.* If autonomy were the issue, discourse ethics would have the instruments to reach an understanding of what ought to be avoided in order to protect personal autonomy. But the very core of the argument is not and cannot be autonomy; after all it is Habermas himself who states:

> Getting used to having human life biotechnologically at the disposal of our contingent preferences cannot help but *change* our normative self-understanding. […] This fact however can count for the heat of the current controversy only as long as belonging to a moral community is still a vital interest.[86]

The instruments to assess that change cannot be moral, because morality, according the Habermas, does not contain an idea of good. Yet, when it comes to judge the desirability of one form of normative self-understanding over another or over no *normative* (not the kind of normativity we have known so far) self-understanding, the categories that one uses are necessarily shaped by some sort of idea of the good, no matter how weak.

Lacking "compelling moral reasons" against positive eugenics, Habermas' way of interrogating himself about the consequences of such practices can only count as proof that the reasoning must be taken on a different dimension. A dimension in which, it is worth remembering, he embarks only in order to protect what he perceives as in danger—the moral community—and in regard of which his radically critical instrument—discourse ethics—has no use.

[86] FHN, pp. 72-73.

Gattungsethik

The conceptual nucleus that brings Habermas to oppose any non-conservative manipulation of the genetic structures of humanity is enclosed in the term *Gattungsethik*: the ethics of the species. Habermas does not use this concept in order to assimilate the status of the embryo to the one of a person. The legal/moral webs of inviolability here play no role. What must be protected is the unavailability of the material substratum of a pre-personal individual life. That is to say that very thing that, according to Habermas, allows each component of the moral community to recognize oneself and the others not simply as human beings but also as beings of equal standing in the community beyond both rational skills, and cultural conceptions of the good.

The danger of losing this common substratum, and, after the domino effect trigged by it, morality itself, is so radical to push Habermas into the space of a minimal ethics. Being unable to justify what is morally wrong about enhancing future human beings, but still suspecting tragic consequences, Habermas states that "where we lack compelling *moral* reasons, we have to let ourselves be guided by signposts set up by the *ethics* of the species".[87] When communicative rationality, organized in the formal structure of discourse ethics, is not able on its own to secure the very existence of morality, we need to look at a weak ethical frame that is able to support or sustain moral discourse.

This kind of relation between species ethics and morality, according to Habermas, is not such to make possible confusions between what is ethical and what is moral. The fields are very well distinct, and anyway the criteria of moral validity keep their formal autonomy. It is morality as a whole that needs "the intuitive self-descriptions that guide our own

[87] FHN, p. 71. Lewens, T. (2012) thinks that the very idea of the species is dangerous.

identification *as human beings".*[88] The species is then the queer mix between a reflective conception of the self and biology that supports and makes morality possible:

> The priority of the just over the good must not blind us to the fact that abstract morality of reason proper to subjects of human rights is itself *sustained* by a prior ethical self-understanding of the species, which is shared by all moral persons.[89]

Habermas has no problems here to admit that discourse ethics is grounded in a form of ethics, namely, species ethics. What turns to be fundamental to the existence of discourse ethics (the functioning is a different matter) is not so much the possibility to recognize each other as rational beings who are equal because they are rational, rather it is essential to be akin in a manner that stands outside rationality, so to speak. The mutual recognition as human and the capacity to recognize a common fate are preliminary to morality.

This passage above could be seen as quite unproblematic: treating others fairly is possible only if first we 'see' each other, and only if we care enough about each other, to put it trivially. However, Habermas has taught us that communicative reason is allegedly so powerful because it is virtually inescapable; here he is telling us instead that, at least potentially, some of us, in the future, could encounter a scenario in which communication is not enough to ensure the existence of the conditions of possibility of a shared self-understanding. This potential scenario generates the necessity to focus on what makes us similar beyond, or before, communicative reason. However, Habermas' identification of what makes us similar with the spices is extremely problematic.[90] In order to understand

[88] Ibid., p. 39.
[89] Ibid.
[90] Engaging directly with the concept of 'species' is not my aim, and even if it were, I would not have the necessary knowledge to sustain such an investigation. Yet, it is important to notice that the concept

why it is so problematic one needs to focus on two related elements on which Habermas draws in his remarks concerning species ethics: recognition, and what he defines as "the intuitive self-descriptions that guide our own identification as human beings", that is the intuitive element which, at least, co-determinates recognition.

If this form of identification, that is the recognition of oneself and the others as akin, is not a product of communicative reason, then it is not linguistically meditated. If it is not linguistically mediated, it cannot be engaged argumentatively, and, therefore, it cannot be critically assessed— so far these are logical steps that Habermas himself lays out for us.

It is Habermas, in fact, who writes that the existence of morality—which for him is identical with discourse ethics—is sustained by a prior self-understanding—that is a form of ethics that we share, or a communality that we feel, pre-reflectively with other members of the human species.

While granting that the interest in being moral cannot be morally normalized, and that therefore, in order to have this sort of impulse, we need to recognize others as similar to us, it is not clear how Habermas intends to justify his assertion that it is, in fact, the species that feeds that vital interest.

Though, while it is not clear how Habermas intends to justify that it is the human species that once intuitively recognized, feeds the form of caring which, in turn, works as condition of existence of discourse ethics, pinning down the problematic aspects of that assertion is difficult. This is because the expression 'human species' which Habermas employs to define our bodily element, instead of suggesting a division between the biological and the rational, seems to allude—particularly for the presence of the

---

Habermas borrows from biology, in that field, more than anything else, is the name of a problem, an ontological one. See the following book, which also provides a survey of the most relevant literature about the problem of the species: D. N. Stamos, *Biological Species, Ontology, and the Metaphysics of Biology*, Lexington Books Ed., New York, 2003.

adjective—to an indissoluble co-presence of the two. Therefore, it seems that all what Habermas is stating is that in order to be moral one needs to be human. However, given that Habermas is interested not in the mere existence of morality, but in the existence of a specific form of morality—discourse ethics—he strategically uses that expression with a very specific meaning.

By using the expression 'human species', he intends to point at an alleged separation between communicative reason, and the human bodily component. In this manner, he has the chance to both present them as two forms of unconditionality, and explain that what is at stake with genetic engineering is not our rationality.

As Taylor explains efficaciously, in a manner I will focus on in Chapter II, the conditions of existence of the "intuitive self-description" cannot abstract from being either an embodied being or a rational one (an in his understanding, a historically placed one). It is, instead because we are both embodied and rational that we can reconstruct the conditions of possibility of the phenomenon of recognition. We can, in fact, start from the very phenomenon of the intuitive self-description and trace back all the logically indispensable elements that have made the intuitive self-description possible: being a self, for instance.

Now, what Habermas is interested in is a specific qualification of that intuition of the self: the idea of co-belonging to the human species. However, his description of how one comes to be a subject makes no convincing argument to the effect that when we do recognize each other as akin, it is the species that happens to be determinant and not, for instance, our common questioning about what is there that makes us akin.

The fact that we actually care or can care about being the same species, in fact, says little about the specific manner both the reciprocal recognition and the consequent caring emerge.

The biological must play an undeniable role in it, as we are embodied beings, yet presenting the species as an element that we can actually sever from the rest of our form of life, and point at as it were the element, rather than one of the elements, that makes recognition possible goes a step further than what Habermas can legitimately maintain.

Habermas' evolutionary account is plausible, and it is historically accurate to state that the body came to play a completely different role in our understanding of our human status compared with the past. For instance, it is no longer the case that one can unproblematically be seen as less than a human on the ground of being a woman, a black person, or an individual unfit to fight, for instance.

Yet, our way to overlook contingent body related differences, and to focus on the underlining elements that we now recognize as our common roots is part and parcel of our whole form of life. We can focus on our shared characteristics (as belonging to the same species) because we care, and not the other way round. And by holding this, one does not need to misrecognize the fact that our bodily component co-determines that very reciprocal recognition. All what I am maintaining is that separating, in a sort of neo-Cartesian fashion, the influence of our species, from the one of our *Lebenswelt* in order to provide a quasi-formal explanation to the caring for each other is a task Habermas is not equipped for.

The understanding of our (potential) caring for each other, and the consequent explanation of what it is that makes the belonging to the same species relevant would require a close engagement with our form of life, that is to say to say, with the so called

modern western ethics—which is the only place where an universal form of caring ever existed.

What Habermas does, by using the already problematic concept of species, is, instead, a form of cherry-picking.

He takes the fact that belonging to the human species matters to us moderns and says that it does so *because* it is the root of our self-intuition, yet he does not actually investigate (a) the origin of that caring as would have required; (b) how self-intuition works, and what all the elements are that makes it possible; and (c) if those elements are so sharply separable that one can point at the one which, once modified, would endanger such a self-intuition.

By carving out from our modern ethical life a concept that is congenial to the existence of discourse ethics, but says nothing about how it comes that belonging to the same species is relevant to us, Habermas only proves that discourse ethics has a chance to work only by relying on a specific part of an equally specific, and substantive ethics. Yet Habermas has no instruments to engage with that kind of ethics.

In this manner discourse ethics demonstrates to be neither a pragmatics, nor a radically critical process. For it could work only with a specific and substantive ethical support (the one in which belonging to the human species is relevant), and yet It does not have the instruments to investigate the motivational elements on which it depends.

Instrumentalization of the inner nature

The fact that Habermas does not employ instruments capable to critically assess whether positive eugenics can or cannot endanger something essential about the present way we

look at ourselves and the rest of the humanity, does not mean that we do not need a way to critically assess it. It is, in fact, possible that we are underestimating the role our body plays in the way we understand each other and ourselves. It may well be the case that, in an ironic way, the objectifying power of functional reason that Habermas has tried to contain and tame within an inescapable reflective intersubjectivity find their better alliance within the plastic body material which lends itself to become totally objectified.

Humanity, uncoupled from rational public discussion, returns in the only way modernity seems to allow it to return: the species. The contingency of the body, humanity as possible calculating object, risks crossing the threshold of the unspeakable, turning humanity into something disputable.

The difference between "being a body and having a body" that is to say the danger to assist to the transformation of the original modern project of entrusting humanity to itself into the self-engineering model, alarms Habermas, but the way he tries to deal with the issue does not meet the standards of his radically critical project.[91]

It is evident that the idea of an ethics of the species introduces a meaningful widening of Habermas' prospective. Unfortunately, this is a sort of opening to (again) a pre-linguistic substratum is the usual way in which Habermas deals with the foundations of discourse ethics: he negates the possibility to engage with that very pre-linguistic substratum on one level only to rescue it on a different one. A level—first religion and now biology—that, according to Habermas, should be less dangerous and less problematic for the functioning and existing of discourse ethics, but that once addressed makes explode the contradictions inherent in his approach.

[91] Mendieta, "Habermas on Cloning: The debate on the Future of the Species", *Philosophy and Social Thinking* 30 no 5-6 (2004): 721-743 and "Communicative Freedom and Genetic Engineering", *Logos* 2, no 1 (Winter 2003)124-140.

Habermas' attempt to revive the modern critical project fails both on the semantic and motivational level, as well as on the ethical level because he reproduces on all levels a variation of the same problem. Discourse ethics is a procedure and as such no internal modifications can occur without radically changing and/or invalidating it. Habermas, then, proceeds to act outside of it, so to speak. He intervenes to manipulate the environment in order to make it the one required by discourse ethics. But these very attempts to provide discourse ethics with the right environment show that it is dependent on external condition that it is not able to assess. And this undermines Habermas' original project of providing us with an instrument that puts us into a position to bring about a post-metaphysical objectivity and enable a radical (self-)critique of reason.

With this chapter I have demonstrated that Habermas' radical critique—discourse ethics—in order to work needs an external aid. This aid, though, cannot be the minimal form of ethics Habermas suggests by drawing on the concept of human species. Such form of ethics is, in fact, unfit for the purpose because it does not engage with, let alone explain, why having the chance to recognize each other as akin is relevant to us. Yet, that kind of inquiring, which equates to an engagement with a substantive form of ethics, is precisely what discourse ethics needs in order to possibly work.

Therefore, in order to further explore the chances radical critique has to work, one would need historico-hermeneutical instruments that Habermas — given his proceduralism — does not have.

It is for this reason that, with the following chapter, I turn on a form of radical critique which, drawing also on substantive sources, has developed also that kind of historico-hermeneutical instruments.

Conjugating the Historical with the Fundamental

Habermas' answer to the question of modernity starts from the question about how to provide a legitimated ground to practical reason. Habermas' approach aims to treasure the distinction between factual truth and normative rightness. The strategic use of such a disjunction, in Habermas' project, aims to make the problematic proliferation of the visions of the good compatible with a (pragmatically) universal justice. The success of such a project would, in the eyes of its author, ensure the emancipation of reason from its idiosyncratic expressions, and, consequently, prepare the terrain for a possible way to tame the distortions of reason itself.

The intrinsically linguistic nature of the foothold Habermas has designed in order to enable such a process of emancipation, places the process itself in a theoretical space that is both practical and epistemic. In other words, the shared (and virtually inescapable) coming to an understanding of what is valid is shaped as a learning process that has an immediate practical consequence: the compelling force of what is understood to be morally valid. This thematization of the just in validity terms allows Habermas to conceive of his pragmatics as a radically critical process which allegedly ensures reason's reasonableness.

Habermas' paradigmatic approach (discourse ethics) to the question of modernity, though, as I have shown with the previous case-studies, faces structural problems. It, in fact, clashes in various ways against the semantic sources it is supposed to validate. In particular it ends up embracing a sort of metaphysical biology that Habermas, given his theoretical instruments, cannot justify.

Taylor, in the paper *The Motivation behind a Procedural Ethics*, notices that if thinkers as Habermas insist on traveling the path opened by a so problematic

understanding of reason it is because they assume "that only a procedural ethics can sustain critical claims" as it appears as this is the only way to allow for "the possibility of an apparent grounding of ultimate ends".[1] Yet starting from the consideration that even the supposed chasm between *is* and *ought* is a historical product that presuppose a certain assumption about the good, Taylor notices that a theoretical construction that avoids to interrogate those very assumptions creates confusion, and ends up failing to achieve its aim. For without critically examining these assumptions, the project falls short of providing a radical self-critique of reason.[2]

The most promising alternative to Habermas' strategy is, in my knowledge, Taylor's own radical critical project. Taylor, in fact, by committing his own work to the notion of epistemic gain—that is the idea that as historically situated agents we are still capable, and licitly so, to discriminate between better and worse ways of being and thinking—takes on a radically critical project that unfolds through an equally radically critical engagement with historical and cultural transformations.

Therefore, the reason I now turn to Taylor is that Taylor's alternative understanding of how an epistemically justified judgment works (at least with respect to practical reason) gives me the chance to investigate whether a theory that embraces a substantive starting point can still be critical in the radical sense, that is to say, by providing an answer to the question of modernity.[3]

---

[1] Taylor, "The Motivation behind a Procedural Ethics", in *Kant & Political Philosophy*, R. Beiner and W.J. Booth eds, Yale University Press, London, 1993, p. 358.
[2] Ibid., p. 338.
[3] I am committed to the reading of Taylor's project as a form of radical critique. Though, if the reader were either not convinced by my reading, or committed to a different interpretation of Taylor, it would be still the case that the issue at stake in this chapter is whether a form of radical critique could be grounded in the sketched way. For, in my view, it denotes the most promising alternative to Habermas — even if this were not Taylor's own position.

According to Taylor the issue of the reasonable working of practical reason cannot be faced in a fruitful manner without recognizing that first and foremost we are the sort of beings for whom things matter, and only insofar they matter can we discriminate among them. It is, in fact, this very ethical orientation that places us in the condition to formulate moral (in a very broad sense) judgments. In addition, as modern version of those beings, that is, as radically reflective version, we are in the peculiar and unprecedented position that we feel the need to radically call into question not only what matters to us, but also the conditions of the judgment of value itself because we are aware that we can be seriously confused about these matters, or misrecognize and mystify them.

What cannot be called into question though is that the conditions of possibility for practical judgment—the hyper-goods in Taylor's terminology—not only are already in place, they also always already orient our judgments. This means that the good has, in Taylor's view, a fundamental enabling function. The simple fact of the orientation to the good is, therefore, for Taylor, the point from where to start the task of thinking the practical side of modern reason.

Starting out from this constitutive orientation of the good, the diagnosis, and the possible cure of the malaises of modernity can be productively faced only if we reach a better understanding of the way this constitutive orientation itself works. And the only way this can be done is, according to Taylor, by studying the identity of the beings for which this issue became visible, and meaningful.

The question about the good is, due to its constitutive function, the other side of the Janus-faced question about who we are. For this reason, according to Taylor, such investigation implies an analysis that unfolds on two interwoven levels: a) the genealogy of the modern identity and the correlated hyper-goods, that is to say, the study of our moral

ontology; and b) the bringing to light of the fundamental trans-historical structures which constitute the formal correlation between the good and the identity.[4]

The complex relation between what has historically emerged, and the fundamental trans-historical structures, which, for Taylor, are completely given in the historical, is the dynamic that needs to be articulated in order to reconstruct the self-comprehension of the modern individual. That is the being that is able to disclose the reflective space necessary to radical critique because it both interrogates and describes herself and her motivations. Said in a more synthetic manner, Taylor suggests investigating self-comprehension as grounding of modern practical reasoning and thus for radical critique.

By choosing his field of investigation, Taylor makes immediately clear that for him critique is not enabled by transcendent criteria of validity that shape and are employed in practical judgments. On the contrary, critique is understood to be made possible by an articulation that discloses, and, to some extent, produces the vital connection between moral ontology (who we have come to be for ourselves) and practical reason through the very practice of bringing to light both what has come to be important to us, and how we relate to our moral sources. Once the articulation of this connection—which, for Taylor, equates to a precondition of self-understanding—is elaborated, it discloses the moral sources already in use, and, in so doing, allows for what he defines an *ad hominem* form of argumentation based on transitions; that is to say a kind of reasoning that, while drawing on the authority of a specific moral background, needs to ensure the epistemic validity of said background through the perspective of the actual agents and the authority of her own judgments.

[4] I am here adopting the terminology—fundamental and trans-historical—that Ricoeur suggested in his analysis of Taylor's, and that the latter explicitly approved in his answer. See the exchange in *Charles Taylor et l'interprétation de l'identité moderne,* edited by Guy Laforest and Philippe de Lara, Paris, Les Presses de l'Université Laval, 1998.

According to Taylor modernity not only compels us to reach the awareness that the critique of our ethico-moral aims can happen only through an evaluative and non-extrinsic language (precisely the articulation of what matters to us), it also provides us with the possibility to engage with practical evaluations from within a historical context such that both the validity of the evaluation are, as far as possible, ensured and the demands imposed by radical critique are fulfilled.

It is against the backdrop of these immanent sources of justification–and given the possibility to study the process of evaluating itself also from a formal and trans-historical perspective (see the fundamental trans-historical structures mentioned above as point b)–that Taylor holds on to the ambition to understand the articulation of what matters to us also "as error-reducing and hence as epistemic gain".[5] He, in fact, thinks that he provides us with a form of interpretative critique of the aims that reason has given to itself over time. Furthermore, he also reckons it to be possible to produce the tools for both a diachronic and synchronic comparison of different historical expressions of practical reason.[6]

Given that Taylor claims that we can look at historically emerged hyper-goods while, at the same time, get into view what Horkheimer once called the objective side of reason, it is fair to argue that Taylor's approach belongs to the family of radically critical endeavors which intend to continue the project of modernity. His approach also amounts to an alternative to approaches which, contextually to the said endeavors, fail when compelled to face the ethico-motivational horizon.

[5] Taylor, SOS, p. 72.

[6] Alessandro Ferrara suggests a strong consonance between this endeavor and what he defines *Reflective Authenticity*—concept to which he devoted his own homonymous book.

In this chapter I explore and assess, drawing mainly on Taylor's *Sources of the Self, Philosophical Arguments*, and selected sections *Philosophical Papers* (I and II), his alternative way to discharge the modern challenge of grounding practical reason.

The way I have chosen to explore Taylor's understanding of self-interpretation as a form of radical critique starts from its micro-level. I disassemble self-understanding into its basic elements and then proceed showing how those elements interact creating Taylor's complex understanding of critique. I proceed in this manner in order to ensure an adequate understanding of his theoretical undertaking to reconstruct the general frame of modernity in such a way that the historical account reinforces and qualifies the non-historical theoretical elements.

In so doing, I focus first on the two dimensions—the formal trans-historical dimension of self-interpretation, and the substantive relational account of the good—that necessarily interact in every act of reflection, but which takes a particular form in modernity. Subsequently, thanks to the case studies provided by Taylor himself, I focus on the manner, the moral agent, through an interpretative dialogue with her identity and history, can i) articulate a ground for her practical judgments, and ii) provide explanations of how this articulation intends to not succumb to relativism. Finally, I examine whether Taylor's description of transitions is actually able to amount to a valid form of epistemic justification.

With respect to Taylor's account of transitions I do not simply argue that his argumentation is not well executed—such a line of critique has also been proffered by Skinner and Nussbaum. What distinguishes my position from theirs is that I argue that making this argument work is structurally impossible for this would require to either take recourse to a naïve conception of history, or to embark on the impossible task to integrate

both new emerged forms of significance and a continuous epistemic growth in the very theoretical construction Taylor has built.

The Trans-historical and the Historical

According to Taylor an act of theoretical reflection is primarily an act of self-interpretation.[7] An individual relates to herself, to her thoughts, and above all to her experiences starting from a place—her identity—that is partly given, and that provides her with the sources of meaningfulness that are required to articulate one's awareness and the contents of said awareness that are involved in the reflective act. That is to say, the individual, in order be capable of self-reflection, at the very least, must reach the comprehension that the one who is experiencing is she. In this theoretical movement of relating to oneself, though, the sources of meaning which contribute to constitute the identity are not the only elements involved; there is also the movement of relating itself. This means, according to Taylor, that for a reflective theoretical act to take place, two dimensions—the trans-historical and the historical—need to be constantly involved. The formal and fundamental structures of self-interpretation which belong to a trans-historical dimension, and a substantive orientation of historical origin must play together in order for a reflective judgment to be produced.

For the purpose of studying the reflective act qua self-comprehension, and to subsequently investigate the critical dimension that this kind of comprehension discloses, it is essential then, for Taylor, to develop his work by constantly focusing on both dimensions and on their interactions.

[7] Taylor, "What is Human Agency?" in *Human Agency and Language, Philosophical Papers I,* Cambridge, Cambridge University Press, 1985, pp. 15-44.

For Taylor, the possibility of formulating a reflective judgment involves both the formal and the substantive dimension. In order to not lose myself in Taylor's overwhelming narrative flux, before engaging in a close examination of the form of radical critique that self-interpretation is supposed to enable, it is thus essential to draw attention to the way Taylor understands those two dimensions, their functioning, and their interplay. In particular, I start the reconstruction of these two dimensions from what Taylor takes to be their respective minimum constituents, that is to say a) the formal and necessary conditions of experience deduced in accordance with forms of argumentation that Taylor defines as 'transcendental', and b) the historically shaped boundary of awareness that forms part of having an experience.

It is, in fact, at this micro-level that the interactions between the formal dimension and the substantive one become manifest, thereby showing that while the relation between the two is an invariant element of the process of self-interpretation, the substantive side opens up the relation itself, and consequently self-comprehension to history and its changes—more specifically to learning processes as Taylor wants to argue.

In the section of the *Philosophical Arguments* that Taylor devotes explicitly to the topic of the validity of transcendental arguments, that is to say, to the study of the formal, trans-historical conditions of the possibility of experience, he describes them as "chains of apodictic indispensability claims which concerns experience and thus have an unchallengeable anchoring".[8] This means that in virtue of these arguments, when a nexus of indispensability is established, its being "indispensable *to* can't be shrugged off".[9] And if this cannot be done, according to Taylor, it is because we cannot coherently formulate

[8] Taylor, "The Validity of Transcendental Arguments" in *Philosophical Arguments*, Cambridge MA, Harvard University Press, 1995, p. 28.
[9] Ibid.

doubts about having experiences.[10] Therefore the categories (coherence and sense of the self for instance) which prove to be indispensable, enjoy the strength of an argument that, even if retrospectively discovered, is an *a priori*.

Yet, exactly because the strength of the logical chain of indispensability is embedded in experience, it is necessary to remember that "transcendental arguments are *arguments*: we need a lot of discourse to establish them because […] we have to articulate the boundary conditions of awareness" that is to say those limit conditions that, if not met, would cause the failure of the awareness of oneself.[11]

Now, when it comes to self-comprehension, which starts first of all by the comprehension of what it means to have an experience, the ultimate boundary conditions of awareness that need to be articulated in order for the chain of apodictic indispensability claims to proceed unencumbered, and I would say meaningfully, from one claim to the other, is the interpreting self. The self is, in fact, the orienting structure that essentially defines, unifies, and limits the field of our perceptions (the ones experienced as such), and the awareness of them.[12] The self, understood in these terms, cannot be the object of an empirical study, as this would involve ontological questions that go way beyond the scope of the transcendental arguments. However, given the fact that the sense of this self is "constitutive of our experience", what can be studied is "what we are inescapably to ourselves".[13] That is to say that the self cannot be studied in itself, but, instead, can be

[10] To be more precise, Taylor does mention also cases in which such kinds of doubt can be formulated, but in that case, the doubt itself rests on the awareness of being engaged in the activity of doubting—a meta-reflection meant to test the limits of doubting without actually undermining the activity itself. See Taylor, "The Validity of Transcendental Arguments" in *Philosophical Arguments*. Cambridge MA, Harvard University Press, 1995, p. 28.
[11] Ibid., p. 31.
[12] Ibid., pp. 28-29.
[13] Ibid., p. 26.

studied by “articulating an insight we have in our own activity”.[14] This insight which in its formal relational structure is trans-historical, for Taylor, is no longer immune to history and therefore to empirical investigation once it is seen from the side of the articulation of the sense of ourselves.

Exploring the self, then, for Taylor, “involves tracing various strands of our modern notion of what it is to be a human agent […] But pursuing this investigation soon shows that you can't get very clear about this without some further understanding of how our pictures of the good have evolved”.[15]

This means that according to Taylor, articulating the question about who we are for ourselves requires us to take constantly into consideration both the trans-historical (the apodictic chains of indispensability), and the historically emerged good(s) as they are constitutive of the identities involved in the processes of self-understanding.

Modernity as the anteriority that places those goods, and consequently the identities of the selves which freely articulate themselves in a constitutive relation to the good, is what makes the relation both possible and problematic—possible because modernity makes the good(s) available so to speak, and problematic because it is the time in which conflictuality, and confusion emerge in relation to the good. It is therefore modernity itself that needs to be mapped-out in order for the connections between the self and the good (the historical) and the relating itself (the trans-historical) operating between the poles (precisely the self and the good) to be fruitfully articulated. It is, in fact, only after these kinds of investigations that an understanding of who we are, and consequently of the conditions that enable our awareness would be available.

[14] Ibid., pp. 26-29.
[15] SOS, p. 3.

Moreover, it is exactly through the study of modernity as the *locus* of historical elements which contribute to constitute ourselves (affirmation of the ordinary life, the expression of the creative inner energies, and above all inwardness) that Taylor is able to clarify and defend the reasons why, according to him, modern critique is unavoidably rooted in self-interpretation.

By establishing that the possibility of a valid form of radical critique is internally linked with self-interpretation and therefore to the bidirectional trajectory that goes from the logical chains of indispensability to the way the individual understands both the good and the self, and the relations between them, Taylor has already set the stage for his next theoretical moves. The interplay between the trans-historical and the historical is, in fact, clearly the key element to explore further. This element is of such importance that Taylor devoted virtually his entire oeuvre to unpack and justify it and to build on it his whole theory of practical reason and its radical self-critique.

It is in virtue of this, that the next sections of this chapter focus on the development of the two dimensions here only mentioned in their basic forms and relations, and on their interactions. Against this backdrop I will then examine the way Taylor put these dimensions to use in his approach to radical critique. In other words, once we are able to appreciate the full complexity of his project, I will be in a position to uncover its problematic nature.

Modernity and the Emergence of Radical Responsibility

The trans-historical and the historical dimensions are both essential to self-comprehension. However modernity is, according to Taylor, what gives to the relation between the two dimensions the specific turn that makes modern individuals radically responsible for who

they are. As moderns, according to Taylor, our relation to the moral sources is, in fact, problematic, and open to questioning in a way it never was before. The dissolution of the "great chain of being" has left the space open for the individual to raise doubts not only about what is right to do, but also about what kind of being it is good to be.

Even if some claims about equality and freedom, for instance, are considered axiomatically valid, when it comes to the moral sources which underpin them, confusion dominates. Taylor explains that "[u]nderneath the agreement on moral standards lies uncertainty and division concerning constitutive goods".[16]

The comprehension of who we are for ourselves, that is to say, our moral ontology as opposed to an unattainable metaphysical knowledge concerning who we are, requires, according to Taylor, first of all, the comprehension of the way in which our deeper motivations came about. It is therefore for Taylor necessary to reconstruct our "genealogy of morals". In other words, he intends to provide an account that is able to (1) "explore the background picture" which both underlies our moral intuitions and functions as context that gives those intuitions their sense, and (2) offer a narration of the origin of modernity in order to being able to present it as an improvement of our epistemic plight.[17]

In so doing, he intends to justify the very shaping of his understanding of critique as self-interpretation and, at same time, to provide it with a horizon which is both significant and authoritative.

Such a theoretical move is supposed to secure Taylor's substantive understanding of practical reasoning on each of the two fronts of validation. That is to say, an historical narration conceived also as epistemic growth—the first front of validation—makes

[16] SOS, p. 498.

[17] In regard to this second point see: "Who knows whether further transformations in the available moral sources may not alter all these issues again out of all recognition? I want to argue that our present predicament represents an epistemic gain, because I think that the alternative moral sources which have opened for us in the past two centuries represent real and important human potentialities". SOS, p. 313.

available to the agent pre-selected valid hyper-goods that she, as self-aware agent—second front of validation—is able to call into question.

This strategy, as I show later on in this chapter, is effective when it comes to providing a foothold for Taylor's criticisms of the one-sided development of the concept of modern reason, but it fails in regard to his attempt to shape a critique which is a) radical, b) substantive, and c) trans-historically and trans-culturally valid. However, before unpacking my critique of Taylor's project I first need to complete my reconstruction with the intention to outline the strongest or most plausible version of it.

Taylor's account of what he reckons to be the long historical roots of modernity goes back to the ancient Greeks. Starting from there, he traces the emergence of three main characteristic modern features.

The first feature is inwardness. It is an inner space which the individual enjoys through reflection, a space to which the individual has a privileged access; it is the locus from which meaning, and, above all, order emanate. The second feature coincides with the emergence of the ordinary life as something of worth, and no longer relegated to the realm of mere necessity. Corresponding to this change, all the passions related to the everyday life—such as dignity—assume a more relevant role. Within modernity, they provide "independent justification" to the sphere of the political. The third and final feature is a picture of nature as expressive force capable to feed the singularity of the modern individual.

Inwardness can be said to be the most relevant characteristic of the "largely unarticulated" modern picture. It concerns a mutation of both the epistemological and cosmological order. Within modernity, reason is no longer part of an ordered universe that includes the individual; on the contrary, it is now seen as residing inside the individual

herself. This sort of relocation makes the subject the new principle of (dis)order, and, therefore, the only ultimate source of justification for knowledge.

By inwardness, in fact, Taylor does not mean the simple conscious relation to oneself, but a sense of interiority that has emerged over centuries, and that subsequently has become the only real space where attempts to ground knowledge can take place.

Giving an account of the process that brought about this first feature, Taylor, in accordance with a well-known *topos* of the history of the western thought, recognizes in Descartes the turning point after which we can actually speak of inwardness in the modern sense. Yet, the beginning of this process precedes Descartes by about ten centuries. It is, in fact, with Plato's *Republic* and in particular with his understanding of the soul as "unique locus where all our different thoughts and feelings occur", that something akin to a unified subject began to be understood as the depository of moral and rational sources.[18] In the Homeric tradition, before Plato's invention of the soul, the parts—even the physical parts—were seen as preponderant over the whole.[19]

Plato's unification of the functions of the soul, together with its being organic to the cosmic order, in Taylor's view, gave rise to an initial form of reflexivity. The access to order, though, for Plato, is possible not in virtue of the soul but thanks to the rationality of the order itself. This means that the turning toward oneself is a way of being illuminated by the order the "soul attends to and feeds on".[20] Consequently, 'turning' can be described neither as a radically reflective movement nor with an inner/outer dichotomy, for there is no actual inner space. Instead, the source available as result of the reflective movement is order itself.[21]

[18] Ibid., p. 118.
[19] Ibid., pp. 117-119.
[20] Ibid., p. 124.
[21] Ibid., pp. 122-124.

It is only with Augustine, according to the narration that Taylor provides, that reflexivity acquires its radical character. In the case of Augustine the soul becomes the source of rationality, yet the access to that rationality is given by the memory of God—the memory of which is guaranteed by God itself. The Christian *agape* requires the individual to turn inward to be accessed, but again, that space is not a space of personal interiority. On the contrary, it is an opening to the alterity which is the only way, according to Augustine, to access a secure rational and moral ground.[22]

In spite of this, for Taylor, it is thanks to Augustine that reflexivity starts to reach a mature formulation, for God is something that ought to be sought in the intimacy of one's own inner space. God, that is to say, something different from the subject itself, is still the source of truth, but the individual needs to deal with herself and her memory as a starting point to ascend to that truth. This means that a certain depth is there to be interrogated, and therefore the reflective turning on oneself can be seen as radical.[23]

At this point, according to Taylor's historical reconstruction, reason was still anchored in, and validated by, its divine source. It is only with Descartes and his "disengagement of reason" that inwardness nearly reaches its present characterization.

The decline of the cosmological order allowed the French philosopher to place the sources of rationality within the individual herself. His understanding of reason as something independent from both the divine and nature made "of inwardness, an inwardness of self-sufficiency, of autonomous powers of ordering by reason".[24] The result of such a "detached engagement" was, according to Taylor, a vision of the universe as

[22] Ibid., p. 169.
[23] Ibid., pp. 134-135.
[24] Ibid., p. 158.

disenchanted and mechanistic—a universe which now easily lends itself to instrumentalization.

The very last element which turned inwardness into something complete was, however, not elaborated by Descartes but by Locke. It is what Taylor defines as the "punctual self".[25] The *res cogitans*, which Descartes separated from the body of the individual, is shrunk by Locke to the point of being identifiable with the agent's pure rational will. In other words, the self, as only depository of reason, is then seen as manifesting itself solely in the agent's consciousness that is an entity that can neither be identified with the material nor the immaterial world. This "perfectly detachable consciousness" is seen by Taylor as the final step toward the cultural domination of a procedural model of reason. "[T]he real self is "extensionless". It is nowhere but in this power to fix things as objects. This power reposes in consciousness".[26]

It is this (for Taylor narrow) understanding of rationality that has led to the moral standard of autonomy and to the subsequent, more recent, conception of practical reason as something valid only when detached from conceptions of the good.[27]

The long path that has led to radical reflexivity, and therefore to the emergence of an inner space as ground for self-mastery, is, according to Taylor, not a purely linear trajectory. "Self-mastery", in fact, evolves in a sort of spurious relation to "self-exploration". The latter is different from the former. While having at its core a radical form of reflexivity, "self-exploration" ultimately values "self-acceptance".[28] This kind of relation to the self is a way to recognize the limits of reason by tempering and counterbalancing those very limits with the passions of the body. Focusing on a tradition that goes from

[25] Ibid., p. 171
[26] Ibid., p. 172.
[27] Ibid., pp. 85-90.
[28] Ibid., p. 179.

Epicurus and Lucretius to Montaigne and Hume, Taylor traces not just a single path for inwardness, but also the shadows it throws, and its internal tensions.[29] In so doing, even before switching to the emergence of what he defines as "the expressive turn", Taylor stresses the presence of an internal complexity of each of the sources of modernity.

One of the most relevant implications of both the slow relocation of epistemic sources, and the differentiation of the ethical ones is, according to Taylor, the emergence of an idea of responsibility for oneself which, in turn, informed the political and ethical life of the moderns.

The emergence of such an idea was accompanied and made even more complex by the affirmation of ordinary life. The path that leads to the emergence of ordinary life is shorter than the one that has led to inwardness, but its origin is more variegated. Taylor draws on a rich palette of sources, which range from literature, to visual arts, to history and philosophy. Though, one of the preponderant factors that led to its establishment is identified by Taylor with the rise of different forms of Protestantism (mainly Puritans) of the seventeenth century.[30]

Within that heterogeneous spiritual tradition—characterized by the refutation of the mediation of the clergy and the hierarchy attached to it—the salvation of the believer is completely entrusted to the single individual and her faith. The "the journey to God" is for the Protestants not a common movement, "each believer rows his or her own boat", Taylor writes.[31] Salvation became, in other words, a personal business. And it is this personal relation between the individual and the deity that opens the way to a reconsideration of

[29] Ibid., pp. 178-184.
[30] Ibid., pp. 215-247.
[31] Ibid., p. 217.

the ordinary life as the only context in which it is possible to recognize the sign of divine grace.

With God understood as incommensurably far away, each and every person stands equally far from the divine. The individuals as well as the religious institutions are seen as similarly lost in the face of the inscrutable nature of God. As a consequence of standing in an equal distance from God, all activities, including the daily ones, were as close as they could be to God.

Far from ratifying the extinction of the religious experience, the distance from divine grace 'simply' transformed faith into something daily and personal. The experience of God was, at that point, a question of personal responsibility for "[w]here a mediated salvation is no longer possible, the personal commitment of the believer becomes all important".[32]

Given this new personal dimension of religion, intentions became the most relevant aspect of moral life. It is in this context that the ordinary life turned into the field of a constant personal battle, the only forum in which the individual can make a constant effort to bringing about some order in a world, a world now at a greater distance from God.

It is, according to Taylor's account, the lack of closeness with God and his divine order which renders it necessary to bring about a new order that in some way approximates the one intended by God.[33] The creation of this order can be brought about exclusively through tireless effort—which thus becomes the only way to salvation.

It is in virtue of this cultural transformation that science and faith present themselves as good allies, or as Taylor put it:

[32] Ibid.
[33] Ibid., p. 228.

> [T]he instrumental stance towards the world has been given a new and important spiritual meaning. Instrumentalization is, in fact, at this point, spiritually essential to preserve both the order that God loves and the believer from getting absorbed by the disorder left from a God that is further away.[34]

Objectification and instrumentalization are now completely congenial to spiritual demands which have as a corollary the shaping and maintaining of a harmony — the sign of divine grace — which is "technological".[35]

In conjunction with the withdrawal of God as sources of order, and the corresponding emergence of a meaningless world, the individual appears rich with internal sources as never before in history. The individual feelings related to ordinary life (in particularly to family life), and to work are the way those sources express themselves. It is the consecration of the passions which orbit around production and reproduction.[36] These passions, as privileged manner to access moral life, therefore acquire a normative power: "we find out what is right at least in part by coming to experience normative sentiments".[37] Normative sentiments from then on have informed not only the moral sphere but also the political one contributing, for instance, to the affirmation of egalitarism.

The affirmation of sentiments as privileged way to access moral life is also linked with, according to Taylor, the third fundamental feature of modernity, that is to say the rise of an "expressivist notion of nature as inner moral source".[38] The Romantic (and post-Romantic) idea of nature sees it as a force—Taylor speaks of an *élan*—which runs through the world and "emerges in our own inner impulses". Such impulses, in virtue of the new place reserved to passions, "are an indispensable part of our access to this force".

[34] Ibid., p. 232.
[35] Ibid., p. 233.
[36] Ibid., p. 211.
[37] Ibid., p. 284.
[38] Ibid., p. x.

Consequently, the only way to reach knowledge of such a moral source "is by articulating what these impulses impel us to".[39]

Following mainly the works of Rousseau and Herder, Taylor describes the originating of the expressive turn as a sort of re-appropriation in modern terms of the need for meaning. The emergence of both a procedural form of reason and a rationalized religion has as impoverished the experiences associated with them.[40] Nature, instead, appears asa meaningful depth, resonating in the personal passions and able to compensate for the loss of meaning.

The moral norm that this expressive force imposes is to "become what one is", that is to follow one's personal inclination: a sort of challenge that requires to live in a manner which is equal to one's own originality.[41]

Predictably enough, in the very moment that self-expression became an independent and essential part of the modern moral life, the modern horizon of significance became problematic. The model of reason derived from Descartes and Locke which has at its core principles of instrumental and detached control of the inner and outer nature, and the consequent affirmation of personal freedom as emancipation from passions and culture, clashes with an idea of nature as moral and aesthetical source that ought to be expressed.

On the one hand, creative imagination and symbolic forces are seen as necessitating expression in order to conform to one's own moral nature, on the other hand, such expressions are the negation of human self-mastery. The fact of autonomy—or depending on the perspective, the fact of expressive creativity—are the source of tensions which are

[39] Ibid., p. 373.
[40] Ibid., p. 376.
[41] Ibid., p. 376.

not external to the moral agent; rather they are constitutive of the way the individual understands herself.

The image of modernity that Taylor depicts appears both rich and problematic. Each of the notions he has described has led to the development of moral sources (hyper-goods) we moderns consider as an advancement concerning both our freedom and knowledge. For example putting the moral and rational sources in the control of the subject has led to the modern concept of moral autonomy; the affirmation of the feelings related to the ordinary life has contributed to the rise of political egalitarianism; while the ideal of an expressive nature has shaped the modern commitment to human self-realization and fulfillment. And above all, the complexity of the modern picture (including its tensions and fractures) has given to the moderns the possibility to directly engage in practices of radical reflexivity—the kind of meta-thinking that allows for calling into question the hyper-goods and therefore even the subject's own identity.

Yet, misrecognition of both the richness and the conflicting nature of our moral sources, according to Taylor, would have, and indeed has, serious consequences at many levels. Ignoring the tensions which engender our consciousness means ignoring our composite needs—needs that were not ignored when the different features of modernity emerged. The tension between self-mastery and self-expression coexists, since the beginning, with the aspiration to a synthesis, as can be already seen, for instance, in the work of Kant. Since then, though, according to Taylor, the tension has been exacerbated by the affirmation of a model of reason which is borrowed from the natural sciences, and which is understood as the only way of validly dealing with each aspect of existence

Such one-sided affirmation deriving from the absence of an actual engagement with the complexity of our motivations, according to Taylor, has brought about epistemic, and social malaises.

According to Taylor' analysis, on the one hand, the concept of self-mastery without any reflection about the good, at a political level, has ended up promoting forms of individualisms that are so strong that they hinder political freedom rather than promoting it. On the other, the subsequent alienation from the public sphere, and the related focusing on the sole aspirations of an authentic self-expression—once deprived of any engagement with the articulation of the horizon of significance in which such expressions are inscribed—risks leaving the space to mere anomie.[42]

On the epistemic level, the identification of reason with a procedural model has given rise to a paradox: the ambition to solve the difficulties in which modern practical reason is caught, have been translated into the formulation of forms of radical critique (on which I will come back later) that instead of facing the origin of the problem have continued to draw on it. Such continuity, according to Taylor, becomes apparent in the fact that claims to knowledge are supposed to be shown as valid even though their foundation is ultimately apodictic. This, in Taylor view hinders the search for a viable way of furthering the critical project for the sake of a notion of radical critique that cannot be underpinned.

The apodictic character of the dominant form of radical critique in conjunction with, and because of, the marginalization of substantive forms of reasoning gave rise to a strong form of skepticism.[43]

[42] Taylor, *The Ethics of Authenticity*, Cambridge, Harvard University Press, 1992.
[43] This argument is virtually ubiquitous in Taylor's oeuvre but it finds its thematization particularly in the paper "Overcoming Epistemology" in *Philosophical Arguments*, Cambridge, MA, Harvard University Press, 1995, pp. 20-33.

In opposition to this unbalanced use and understanding of the modern sources, Taylor suggests that we should be conscious that the whole picture of modernity functions as a historical boundary of the modern awareness, and that, consequently, the issue of practical reason needs to be addressed starting from the modern subject's identity and its orientating function. The form of critique that Taylor is after is supposed to do justice to each of the characteristic features of the modern picture:

> What reflection in this direction would entail is already fairly well known. It involves, first, conceiving reason differently, as including — alongside the familiar forms of the Enlightenment — a new department, whose excellence consists in our being able to articulate the background of our lives perspicuously. […] And along with this goes a conception of critical reasoning, of special relevance for moral thinking, that focuses on the nature of transitions[44]

It, therefore, can renounce neither to the substantial elements of reason, nor to the modern demands for validations.

Engaging with both elements, for Taylor, has meant to articulate the hyper-goods in such a manner to make them available to us for further scrutinizing and comparative evaluation. And given the fact that articulating "the background of our lives perspicuously" entails the questioning of our understanding of our deeper motivations, a form of critique which tries to keep together the horizon of significance provided by history and the ability of the modern agent to question that very horizon, needs to be rooted in self-interpretation. Self-interpretation which is saved from arbitrariness by the understanding of modernity as presented in *Sources of the Self*. That is to say that modernity is a general

[44] Taylor, "Overcoming Epistemology", in *Philosophical Arguments*, Cambridge, MA, Harvard University Press, 1995, p. 15.

horizon of significance which is also authoritative given both (a) the presence of the characteristic modern hyper-goods that we can recognize as advances, also epistemologically speaking, in comparison with pre-modern (accounts of) hyper-goods, and (b) the ability to incorporate the possibility to call into question even those hyper-goods.

The fact that Taylor uses history also as source of epistemic justification has been the main object of criticisms.[45] However, anticipating the criticisms, Taylor has clarified that the very endeavor of historical reconstruction of the sources of the modern self does not claim for itself an explanatory power if by explanatory one understands a form of "diachronic causation". His is a "historical account" focused on the will to comprehend, through the articulation of the visions of the good present in our time, the motivations (the hyper-goods) that sustain our self-understanding. By implication, the historical process is presented by Taylor as something that is retrospectively attributed. But this does not change, according to Taylor, the reasonableness of his account. It, instead, gives us the chance to see how the complex ways in which we are moved by such motivations, can be interpreted as a sign of an improvement when compared with the past.[46]

As my reconstruction of the second front of the justification – that is, of the way in which individuals articulate their own motivations – will render apparent, Taylor's clarifications concerning the nature of his historical account of modernity are insufficient *vis-à-vis* his ambitions of having a form of critique that is substantive and epistemologically valid at both the trans-historical and trans-cultural level.

[45] About the criticism of having written a too intellectualistic, selective, and simplified history see among the several: Martha Nussbaum (Nussbaum, 1990, pp. 32-34), Stephen Clark (1991, pp. 201-205), Quentin Skinner (1991), Judith Shklar (1991), Holmer Steinfath (1991, pp. 106-107)

The Articulation of Modern Critique

The intuition that the experience of oneself is always already oriented, that is to say, that this reflective act involves the relating to that orientation which the self is, has been analytically divided by Taylor into what are, according to him, its two constitutive elements: i) the awareness of relating to the self, plus all of the other fundamental trans-historical categories one can retrospectively discover by analyzing the relating itself, and that which is indispensable for self- awareness to exist; and ii) the understanding of notions of self and the good as forming the historical boundary of that relation.

In section two of this chapter, I have focused on the latter by giving an account of how, according to Taylor, our moral ontology has come to be what we understand it to be. In other words, I have roughly reconstructed Taylor's historical analysis of the modern self, and its moral sources.

This historical analysis, according to Taylor, shows the emergence of ideas and ethical visions by which we understand the modern identity, and, at the same time, it also allows us to reflect on the philosophical consequences those historical elements have on the act of self-reflection.

Inwardness, the idea of being able to master and express one's own self, as well as the affirmation of the ordinary life and all the emotions related to it, all contribute to a composite picture from which two elements stand out: a) the inarticulacy of modern moral sources, and b) the consequent radical responsibility for who one is which translates into the responsibility to articulate exactly those moral sources.

In this section, I retrace precisely the analysis of that responsibility as explored by Taylor in the essay "What is Human Agency?"[47] Analyzing mainly that specific work, I provide an account of responsibility no longer seen from a merely historical point of view, but as forming part of the essential practice of articulating our moral sources. Such an analysis allows me to give a more detailed explanation of the meaning of such an articulating means and how it results from the interplay between the trans-historical and the historical dimensions engaged in self-articulation.

According to Taylor, the manner humans engage in the reflective practice of self-evaluation implies, and has always implied two hierarchically connected orders of desire.[48] When an individual interrogates herself in order to deliberate, in fact, the different desires she experiences are evaluated on the ground of her higher desires. That is to say, that as agent, I am in the position to avoid eating too many of the sweet things I would like to eat now, because I may think I have the chance to eat them later, when I will also have the possibility to dine in front of a nice landscape.

The formation of a desire about my desire gives me the chance to distance myself and evaluate my immediate craving for sweet things, and decide that I prefer to postpone the fulfilling of my craving, or to second it, or whatever other decision I will make once I have clear in mind my higher desires.

Among the several options, I might decide to restrain myself because, for me, it is important to be the kind of person able to moderate the excesses that my first-order desires present me with. And this too would be a second-order desire, but according to

[47] Taylor, "What is Human Agency?" in *Human Agency and Language: Philosophical Papers I*, Cambridge MA, Cambridge University Press, 1985.
[48] Ibid., p. 16.

Taylor, it needs to be differentiated from the one mentioned in the first part of this example.

In fact, differently from Frankfurt, in the wake of whom Taylor develops this reflection, the latter holds that an account which sees the formulation of second-order desire as the main characteristic of human agency is in need of yet further qualification.[49] For Taylor a complete account of the ways human agents reflectively deal with their desires still can, and actually must, discriminate between strong and weak evaluations.[50]

This additional differentiation is necessary to draw attention to what Taylor defines as 'qualitative evaluation' of desires and motives. Taylor discriminates between weak and strong evaluations—in the case of the former the agent is concerned mainly, but not only, with the outcomes of her decisions, while in the latter case what is at stake is "the qualitative *worth* of different desires".[51]

The explicit presence of the reference to qualitative differentiations only within the definition of strong evaluations does not suggest that in deliberations in which what counts is the output, the desires involved are seen as homogenous by the agent.[52] Qualitative differentiations, of some form, also form part of weak evaluations. But such differentiations are not concerned, according to Taylor, with the value of the desires themselves.

A person can—in accordance with Taylor's criteria for weak evaluation—deliberate between following her desire to spend her holiday in the sunny south rather than in the north on the ground of what she reckons as a more valuable option. The relaxation the sunny south offers can be regarded as more valuable than a potentially more exciting trip to the north, for instance. The term 'valuable' is used here with a weak meaning, since it is

[49] H. Frankfurt, *Freedom of Will and the Concept of a Person*, «Journal of Philosophy», 67 (1971), 1, p. 5-20.
[50] Taylor, "What is Human Agency?" in *Human Agency and Language: Philosophical Papers I*, Cambridge MA, Cambridge University Press, 1985, pp. 16-17.
[51] Ibid., p. 16.
[52] Ibid.

not the case that holidays in the north are in themselves more valuable than the ones in the south, or vice versa. The desirability of the object, as Taylor explains, is justified by the mere fact that it is desired at that given moment.[53]

Regarding the strong evaluations, things are substantially different. This kind of evaluation is not based on contingent preferences; rather, it is "part of a mode of life".[54] It requires the agent to be engaged in a kind of self-reflection that culminates in the formulation of a judgment concerning the intrinsic superiority or inferiority of the evaluated option. And such an evaluation can be made possible only by interrogating what Taylor, using a metaphor, calls the 'depth' of the agent. To put it differently, such a choice is the result of wondering "about what our motivation really is, how we should truly characterize the meaning things have for us".[55]

When an individual engages in this kind of evaluation what is at stake, then, is her ability to see herself as an autonomous agent who has a specific identity.[56] An identity that, in some sense, is produced by the strong evaluations. That is to say, the identity of the agent, in Taylor's understanding, emerges, in a non-causal fashion, through the articulation of those strong evaluations, for the way an individual understands and experiences herself depends on exactly the articulation of those kinds of motivations. In Taylor's own words:

> These articulations are not simply descriptions, if we mean by this characterizations of a fully independent object, that is, an object that is altered neither in what it is nor in the degree or manner of its evidence to us by the description. […] On the contrary, articulations are attempts to formulate what is initially inchoate, or confused, or badly formulated. But

[53] Ibid.
[54] Ibid., p. 19.
[55] Ibid., pp. 22-23.
[56] Ibid., p. 34.

> this kind of formation or reformulation does not leave its object unchanged.[57]

This possibility to use language to engage with oneself in a manner that allows us, "to characterize the desires as higher and lower", places the individual who is occupied with self-articulation at the center of her experience of agency *qua* agency.[58] In other words, the connection between self-interpretation and experience is constitutive, for the articulation of the strong evaluations—or hyper-goods according to the evolution of Taylor's terminology—partly shapes experience.[59]

Going back to the example mentioned at the beginning of this section, if I am not the kind of person who values self-control, eating too many sweet things will, perhaps, be simply a pleasant experience. Though, if I am that kind of person, my conception of self-control as something good shapes the experience of myself as a dissatisfied individual who has not been able to live up to her own standards. While reflecting on my experience of dissatisfaction, I might even start to think that the way I am feeling is excessive, that perhaps this kind of self-control is damaging other important aspects of my life. Articulacy enables the agent to interrogate what is considered fundamental. Hence, the experience of who I am forms itself while I process and understand my options, values, and relation to them.

[57] Ibid., p. 36.
[58] Ibid., p. 26.
[59] The themes central to Taylor's work are rather constant, but the terminology he employs varies accordingly to both the evolution of its thought and the angle from which he decides to tackle those themes. In the Sources of the Self, where the focus of the analysis was the historical evolution of the conceptions of the good, he defines the pillars of the modern identity no longer as higher motivation but as hyper-goods (see pp 63.73).

The articulation brings to light what I value and gives shape, through a pre-existing language, to both the relation I have with myself, and my self-experience. It is, for Taylor, the possibility of this kind of articulacy which makes the agent responsible for herself.

Yet this general picture is modified in Modernity. The modern agent, in fact, no longer enjoys any kind of "unreflective encapsulation" as those who could count on an ordered and meaningful universe did.[60] The modern conceptual equipment, the clashes internal to it, and the idea of being the ultimate source of meaning and validation, allow the agent to call into question both herself and her hyper-goods.

In this context, articulating one's deepest motives gives to the expression "being responsible for oneself" a more radical meaning. The process of self-interpretation no longer equates simply to the question about which interpretation is "more faithful to reality".[61] A "conflict of self-interpretations", for a modern agent goes a step further as it were. It is able to distort or change the "reality concerned" in a radical manner.[62]

Modernity, in fact, in some occasions, opens up the direct possibility for hyper-goods and experience to influence each other in ways that, according to Taylor, were virtually impossible before. The articulation of an intuition can now modify the way the agent understands herself, not just by discarding one pre-existing hyper-good for another pre-existing hyper-good, but in the sense of modifying the hyper-good itself. The modern agent, by articulating her strong motivations and higher desires, is no longer simply giving to herself "the possibility of a plurality of visions which there was not before"; she has the possibility to question the meaning and the validity of each of those visions. This makes responsibility more radical than ever before.

---

[60] Taylor, "Explanation and Practical Reason" in *Philosophical Arguments*, Cambridge, MA: Harvard University Press, 1995, p. 59.
[61] Ibid., p. 22.
[62] Ibid., p. 22.

This, for Taylor, is very far from meaning that modernity is the reign of arbitrariness in which everything is possible to the point of annihilating the meaning of responsibility and critique. Both responsibility and radical critique are understood as being meaningful because the formation of second-order desires depends on the pre-existence of both the hyper-goods and the language the individual uses in order to articulate the good.[63]

It is, in fact, for Taylor, exactly the preexistence of both the language with which the agent articulates the hyper-goods, and the wider modern horizon of significance (that is the rich, conflicting, and widely unarticulated cluster of the modern social imaginaries) from which the hyper-goods emerge to make it so.[64] The object of Taylor's investigations, in fact, is a radical form of articulation and not a radical choice.[65]

In explicit contraposition to the Sartre of "Existentialism is a Humanism", for Taylor being responsible for oneself does not mean that the individual is able to choose herself. For "this total self-possession would in fact be the most total self-loss".[66] That is to say, either one chooses oneself in a radical manner and therefore for no reason, or the choice about oneself is made on the ground of some reasons.

In the first case, it would be a sort of pure decision taken from an "extensionless point"; in other words, a will without identity would make this first scenario fade into the second one. For, as soon as a choice is made on the ground of reasons and/or values, it means that there is already a horizon of evaluation that has not been originally chosen.[67] This proves, according to Taylor, that when it comes to the shaping of identity, when one

[63] Ibid., p. 37.
[64] The expression 'modern social imaginaries' is an expression Taylor will elaborate only later on with the homonymous book. I use it to expand on what in the concerned paper are only theoretical gestures summed up by the reference to "the rich language" and "the inarticulate limit" from which the articulations originate. See p. 40
[65] Ibid., p. 37.
[66] Ibid., p. 35.
[67] Ibid., p. 29.

thinks of who we are also as the result of our choices and interpretations, it would be inconsistent to think of them as the result of a radical choice, as it were, a picking *ex nihilo*.

The radicalism of responsibility regarding one's own identity, instead, must be understood exactly in terms of strong evaluations "[f]or a radical choice *between* evaluations is quite conceivable, but not a radical choice *of* such evaluations".[68]

If this limited existential prerogative, according to Taylor, does not lead itself to the belittling or even the negation of responsibility concerning our identity, but rather discloses the possibility of its radical character, it is because the reflection about the conditions of possibility for choosing and judging does not casually determine the way individuals actually choose, judge, and experience those actions.[69]

The constitutive relation between articulation and experiences—even if does not make of identity an act of radical choice—leaves in the hands of the modern individual the freedom to shape her being as faithful to herself.

For Taylor this means that not only the understanding of what is important to the individual creates specific conditions for the experience the individual has of herself, but also that new and unarticulated experiences can push the individual to call into question the higher desire(s). The activity of articulating the new experience, in turn, can lead to modifications of essential parts of the way the individual understands herself.[70] The agent is then potentially able to re-consider herself:

[68] Ibid.
[69] Ibid., p. 40.
[70] Ibid., p. 39.

> This kind of re-evaluation will be radical […] in the sense that our looking again can be so undertaken that in principle no formulations are considered unrevisable.[71]

This is what would happen when a doubt surprises the agent about her higher motivations not in the form of "Am I really being just?" but rather about the hyper-good itself. To give an example, this is the scenario in which I have been a person who thought that her life was inspired by a deep sense of justice, and then one day I experience something that makes me question my previous account of what justice is. Once embarked on that kind of meta-questioning, the agent is caught in a re-evaluating process that casts doubts on every formulation of justice.

According to Taylor, the solution to this kind of dilemma cannot come from hypothetical external criteria, rather it starts looking into the particular cases that have provoked the doubts while being "in a stance of openness".[72] Again, this does not mean that whatever answer the agent will come up with is acceptable, "but rather that what takes the place of the yardstick is my deepest unstructured sense of what is important, which is as yet inchoate and which I am trying to bring to definition".

This rather sketchy suggestion about how to face dilemmas of such importance is, what Taylor draws on in order to arrive at a full-fledged conception of the way radical critique as self-interpretation can be "of especial relevance for moral thinking, that focuses on the nature of transitions".[73]

---

[71] Ibid., p. 40.
[72] Ibid., p. 41.
[73] Taylor, "Overcoming Epistemology", in *Philosophical Arguments*, Cambridge, MA, Harvard University Press, 1995, p. 15. Here the full quotation: "What reflection in this direction would entail is already fairly known. It involves, first, conceiving reason differently, as including—alongside the familiar forms of the Enlightenment — a new department, whose excellence consists in our being able to articulate the background of our lives perspicuously. We can use the word "disclosure" for this, following Heidegger. And along with this goes a

Radical Critique

Finding a way to understand how the play between the historical and the trans-historical brings about the reflective space that, in turn, allows articulating self-comprehension has its own theoretical value, but in the case of Taylor, it has to be seen, primarily, as forming part of a wider project. Its study "is meant to serve as the starting point for a renewed understanding of modernity".[74] And while "coming to comprehend the momentous transformations of our culture and society over the last three or four centuries and getting these somehow in focus" is an essential part of this undertaking which, according to Taylor himself associates his work to "the works of major contemporary thinkers such as Foucault, Habermas, and MacIntyre", the philosophical route he has undertaken diverges from each of them.[75]

Taylor's specific understanding of both the historical and the trans-historical elements involved in the play of modernity has brought about his conviction that the cure to modern malaises, moral skepticism being the most important one, needs to pass through the restoring of practical reason to its substantial character. This conviction, which mainly draws on the inescapable orientating function of the good, once combined with the radical form of reflexivity which characterizes modernity, translates into the shaping of a form of critique deeply anchored in self-interpretation.

As shown in sections three of this chapter, Taylor has explained that an embodied agent orients herself in the world starting from a certain conception of the good. Such

---

conception of critical reasoning, of special relevance for moral thinking, that focuses on the nature of transitions in our thought, of which "immanent critique" is only the best known example."

[74] SOS, p. ix

[75] Ibid.

conception is not arbitrary; it is embedded in the horizon of significance which constitutes the historical conditions of the agent's ability to formulate judgments.

Now, trying to bring together the theoretical elements Taylor provides and that I have so far reconstructed in a separate manner. In order to formulate a judgment in a self-aware manner, at least, the agent must be aware that the one who is formulating the judgment is she. This essential formal minimal requirement, that is to say, the awareness that there is an *I* who is formulating a judgment and that that *I* coincides with who is aware of the whole process, immediately encroaches on a historical terrain.

Having the phenomenologically justified awareness of being that *I*—the minimal requirement of every self-aware act—does not place the agent in the condition to know herself as she is in herself through experience. Instead, the experience of herself places this agent within the possibility of knowing her identity that is who she is for herself. In other words, given that the pure ontological questions are precluded to the agent, such agent can still come to know what Taylor defines her moral ontology.

The agent has the possibility, according to him, to know who she is for herself through the mediation of what it means to be a subject in her time, and, above all, through what she considers her fundamental motivations. Said still with other words, since the sense of the self can be turned into knowledge the self is aware of—I am the *I* who is having an experience—only through at least a vague knowledge of the identity involved, then the formal process of self-awareness is necessary bounded by the historical conceptions of the self—the ones which contribute to constitute the actual identity.

Because the interaction between the historical and the trans-historical is constitutive of identity for Taylor; and because this formation of identity is inextricably bound to the conception of the good—which develops over time as a result of the same

historical/trans-historical interactions—the interaction also forms part of the formulation of practical judgments.

The presence of the fundamental orientation given by the historically shaped conception of the good has as consequence the formation of compelling second order desires which give to the agent a manner to formulate her judgment of value.

In unpacking the dynamic according to which the agent articulates (and in this manner also partly shapes) her higher desires, Taylor shows that the articulacy of the hyper-goods not only provides substantial standards for practical judgment, it is also constitutive of the agent's moral experience.

The reflective ability of the agent to deliberate and therefore to autonomously discriminate among the hyper-goods that appear to her as already relevant, together with the fact that the endorsed ones are constitutive of her experience as agent, according to Taylor, make of such agent a person responsible for who she is.

The study of both the historical and trans-historical conditions of possibility of the process that brings about the reflective formulations of the hyper-goods displays the interconnections which keep together identity, responsibility, and practical judgment. The displaying of the fundamental interactions among these elements exhibits an initial plausibility and, for Taylor, the necessity of a concept of (radical) critique qua self-interpretation.

While both formal and historical studies of the formulation of a practical judgment suggest the necessity of such a kind of critique, historical studies, that is to say, Taylor's own study of modernity, makes also evident that modern critique faces complex validity issues.

In his own attempt to articulate both modern hyper-goods, and the related conceptions of the self, Taylor shows that because of both the fragmentation and conflicting character of the horizon of significance, and the correlated affirmation of a radical form of reflectivity, the only ultimate authority concerning the validity of knowledge is the agent's own rationality.

This kind of authority, within the general picture of modernity, makes of both responsibility and critique something radical. The newly acquired possibility of the agent to call into question even her most fundamental moral sources and therefore her own identity raises a question concerning what justifies critique.

The consequence of such a questioning of validity is, according to Taylor, a deep skepticism applied to, above all, those fields, as the practical one, where instrumental principles are of no use. However "the widespread belief that […] moral differences can't be arbitrated by reason", according to Taylor, is a mistake. This mistake derives from a mix of factors, and in particular from a fundamental misunderstanding concerning the proper functioning of practical reason, of critique, and of the role strong evaluations have in relation to them and to modernity.[76]

According to Taylor, the widespread subjectivism which derives from modern skepticism is also favored by actual conflicting experiences concerning moral disputes. There are, in fact, cases in which the gap between cultures amounts to "a mutual incomprehensibility of human moralities".[77]

This enormous problem, though, in Taylor's understanding of practical reason, is not equal to an architectonical problem of practical reason itself. He writes:

[76] Taylor, "Explanation and Practical Reason" in *Philosophical Arguments*, Cambridge, MA: Harvard University Press, 1995, p. 34.
[77] Ibid., p. 55.

> And yet, I want to argue that the considerations above on practical argument show that we shouldn't give up on reason too early. We don't need to be so intimidated by distance and incomprehensibility that we take them as sufficient grounds to adopt relativism. There are resources in argument […] We have to try and see.[78]

Taylor's own way to try and see starts from the analysis of what he reckons is the deepest root of the mentioned modern moral skepticism. It is, in fact, through the close examination of the origin of the skepticism concerning the possibility to arbitrate moral differences by the means of reason that Taylor arrives at his attempt to elaborate an alternative strategy. His alternative strategy, as I will show in the following pages, focuses on the transition between two competing views. It is allegedly capable to "give a convincing narrative account of the passage from the first to the second as an advance in knowledge, a step from a less good to a better understanding of the phenomena in question". That is to say, Taylor's strategy would be allegedly capable to provide the justification for seeing the transitions as constituting an epistemic gain.[79]

According to Taylor, what has contributed the most to this "climate" are "rather deep metaphysical assumptions" In correspondence to the rise of modern physics, a precise model of reason has conquered a dominant position in the modern common understanding. In Taylor's exemplification of this model:

> Our knowledge claims are to be checked, to be assessed as fully and responsibly as they can be, by breaking them down and identifying their ultimate foundations, as

[78] Ibid., p. 55. And on a similar note on pp. 58-59 of the pertinent paper: "it may be virtually impossible, and certainly hazardous, to try to argue people over it. But what does this say about the limitations of reason? Nothing, I would argue […] Even the most exotic differences don't therefore put paid to a role for reason."
[79] Ibid. p. 59, et passim.

> distinct from the chain of inferences which build from these toward our original unreflecting beliefs.[80]

Under the influence of both Descartes' and Locke's work which has contributed to thematize and, in a way, legitimize such a method, it has ended up being completely identified with reason. In other words, moral disputes are seen as something that can be rationally solved only in virtue of principles able to remain valid regardless of the actual dispute, as the process according to which one decides the rightness of moral judgments could cut off from the commitments at stake in the single token of practical reasoning.

Yet, for Taylor, "this identification of the demands of critical morality with the procedural understanding of reason […] is deeply mistaken".[81] Moreover, this identification has marginalized a more promising way of reasoning which does not deal with knowledge starting from a procedural point of view, but rather focuses on a sort of local standpoint. This alternative approach "starts off on the basis that my opponent already shares at least some of the fundamental dispositions"—what he defines as *ad hominem* reasoning.[82]

The way Taylor intends to correct the identification of reason with its procedural model, goes hand in hand with his project of showing that radical critique as self-interpretation is a viable option. For reducing the grip of this procedural model, which ends up underpinning modern skepticism, helps restoring the understanding of practical reason to its substantive character.[83]

The implementation of this task means, first of all, capitalizing the force of the argument concerning the constitutive role of the good, and the related practice of strong evaluations. This implies showing that the interrogation of the validity of the moral

[80] Ibid., p. 40.
[81] Ibid., p. 41.
[82] Ibid., p. 36.
[83] Ibid., p. 42.

argumentation needs to be carried on, while putting a stop to the "tendency to think out the question of what something *is* in terms of the question how it is *known"*.[84]

According to Taylor, in fact, this tendency of focusing exclusively on the manner a piece of knowledge is reached rather than on what that knowledge is, translates into the idea that moral justification is always "(a) effected on the basis of criteria, (b) judging between fully explicit positions, and (c) yielding in the first instance absolute judgment of adequacy or inadequacy, and comparative assessments only mediately from these".[85] However, complex moral deliberations cannot always be brought about—effected—on the ground of criteria thus understood. For it can be the case that either (i) the criteria are not fully articulated, and, therefore, they cannot be engaged with in a reflective manner (yet), or (ii) there are no available common criteria that can bring an end to problematic moral disputes.

For instance, it might be the case that (i) I *prima facie* judge euthanasia as always admissible because I see it as concerning merely the autonomy of the person involved. Then, while trying to make fully clear to myself or to an interlocutor what I mean by autonomy and why I am confining my reasoning to that principle, it can happen that inchoate criteria which I am either bracketing or applying, once they are fully brought to light, show that I am grounding my reasoning in some "special pleading" which undermines the validity of the whole reasoning. This increase in my self-understanding would then bring me either to change my initial judgment, or to keep the judgment but now justifying it in a way that takes into account the (reconsidered) principles I now know I am using.

A further option deriving from my engaging with my motivations can be (ii) a dispute regarding no longer the clarity, articulacy, and consistency of the criteria in use, but

[84] Ibid., p. 34.
[85] Ibid., p. 60.

the validity of the criteria themselves. In this scenario, despite the internal coherence of my reasoning, I have to engage with the very foundation of my process of justification, therefore I cannot appeal to a criterion x that is able to straighten the question out.

When it comes to questions of this kind, according to Taylor, the comparison between opposing moral visions is very similar to the comparisons between entire scientific paradigms.[86] For the rationality of the scientific paradigms as wholes is neither something that can be judged by the criteria ingrained into the paradigms themselves, nor through "externally defined standards" as they are nonexistent in that case.[87]

On a closer analysis then, not even the natural sciences always have at their disposal shared criteria of validity. This very observation, for Taylor, far from endorsing an even stronger skepticism, is the beginning of the right argumentative path to reject it.[88]

History of science, in fact, for Taylor, can be hardly seen as a random cluster of paradigms. Looking at the real historical transitions, it would be implausible to describe the Galilean physics as less capable to make sense of the phenomena concerned compared to what the Aristotelian physics was capable of.

The succession between these paradigms, instead, traces the trajectory of an epistemic gain, or in Taylor's own words, "an advance in knowledge, a step from a less good to a better understanding of the phenomena in question".[89]

Therefore, following the similitude that to him appears as encouraging, Taylor explains that when dealing with the difficulties concerning the possibility to measure the value of one paradigm against the other, what initially looks like a "standoff", once one

[86] Ibid. p. 47.
[87] Ibid., p. 42. Taylor is here explicitly drawing on A. MacIntyre's argument about the possibility of rationally comparing scientific paradigms even in absence of criteria understood as standard externally defined. See A. MacIntyre, *Epistemological Crisis, Dramatic Narrative, and the Philosophy of Science*, «The Monist», 60, 1977 pp. 453-472.
[88] Ibid., pp. 46-47.
[89] Ibid., p. 42

stops pretending that it is possible to judge from a god-like position, and focuses on the actual disputes, can reveal itself as something "arbitrable".[90] In other words:

> we can sometimes arbitrate between positions by portraying *transitions* as gains or losses, even where what we normally understand as decision through criteria — qua externally defined standards — is impossible.[91]

This alternative grounding of rational argumentation focuses on the ability of those who articulate the judgment to provide a *comparatively* better understanding not simply of the 'the facts' involved but also, and more importantly, it focuses on "the ability of each to make sense of itself and the other in explaining these facts".[92] That is to say the analysis needs to zoom out from 'the facts' that have caused the crisis of the usual way of judging in order to focus on the pre-understanding of those facts in relation to the crisis. In yet other words, this way of reasoning requires a focus on the manner we interact with the world, and on the ways and the reasons that motivate us to interrogate it. It is, in short, an interrogation of our form of life. Accordingly, what is at stake, both in the case of the apparently irresolvable moral disputes, and in the shifting from one paradigm to another, are the pillars of the existence of the points of view involved.

Quoting Mill, Taylor reinforces what is a key element of his practical vision that is to say "questions of ultimate ends are not amenable to direct proof". And yet as Mill himself notices "considerations may be presented capable of determining the intellect either to give or to withhold its assent to the doctrine […] and this is the equivalent to proof". [93]

---

[90] Ibid., p. 47.
[91] Ibid.
[92] Ibid., p. 44.
[93] J. S. Mill, *Utilitarianism*, Hackett, Indianapolis, 1979, pp. 4-5. The quotation appears in Taylor p. 36

This brings us to Taylor's initial observation concerning procedural reason, its being recognized as the exclusive way of critical thinking, and the putting aside of a different but fruitful argumentative method.

In absence of criteria to which an agent could appeal, and in need to produce pertinent considerations, the agent cannot but start her attempt to formulate relevant considerations by looking at where exactly her reasoning got interrupted. That is to say, she needs to articulate her fundamental orientation in relation to the *explanandum.* This, which is a process of self-understanding, includes questions about why she is no longer able to make sense of the phenomenon, about the meaning the phenomenon was suppose to have, and whether it is the case that reasons she was marginalizing make more sense both of the phenomenon in question and of the reasons why, within her own perspective, the phenomenon is problematic, and so on.

Such an approach, in the case of natural sciences, can be theoretically unproblematic for the reasoning in that field has a functional manner to measure itself and the potentially new explanations provided. According to Taylor, practical knowledge is "intrinsically linked with increased ability to effect our purposes" at the point that "an extension of our practical capacities is therefore a reliable criterion of increasing knowledge".[94] Yet this very same approach in the field of practical judgment can be seen as incurring the problem of deciding which "extension of practical capacities" one ought to endorse.

The reasoning *ad hominem,* according to Taylor*,* could risk, within the moral field, to crash against the same problem Mill encountered: the natural fallacy. The British

[94] Taylor, "Explanation and Practical Reason" in Philosophical Arguments, Cambridge, MA: Harvard University Press, 1995, p. 48.

philosopher ended up justifying the desirability of happiness through the fact that happiness is desired. Mill, thus, confused the descriptive level and the normative one.

Yet, Taylor reckons as incorrect the hypothesis of associating also his own work with the same fallacious form of reasoning. He explains that his suggestion of using this *ad hominem* reasoning is backed by his distinction between strong and weak evaluations. That is to say, the distinction between those evaluations which are formed on the ground of *de facto* preferences (weak evaluations), and the strong ones, that is to say the evaluations formulated in relation to the (historically selected) agent's hyper-goods. Given this distinction, the gap between the descriptive (the articulation) and the normative shrinks without invalidating the reasoning, in his view.

Taylor is completely aware that by bounding the validity of the practical judgments to self-comprehension, he is committing himself to a sort of riddle, but in his own view he is also convinced to have provided his critique with the necessary theoretical balance and counterbalance in order to navigate safely through it.

The "biographical narrative" individuals need to engage with in order to articulate the moral sources which underpin who they are, and therefore to shed some light on their conception of the good is, at no level, a transparent and easy proceeding. And this is exactly because, as indicated in section three, "these articulations are not simply descriptions, if we mean by this characterizations of a fully independent object, that is, an object which is altered neither in what it is, nor in the degree or manner of its evidence to us by the description".[95] The articulation, in other words, "does not leave the object unchanged".[96]

[95] Taylor, "What is Human Agency?" In *Human Agency and Language: Philosophical Papers 1*. Cambridge, Cambridge University Press, 1985, p. 36.
[96] Ibid.

Yet, Taylor explains that from this it does not follow that articulations are arbitrary. We can still, according to him, speak of wrong or truthful articulations. This is possible because when individuals, with their descriptions, "try to be faithful to something", even while being unable to reach an ultimate knowledge, they move within historically given coordinates—the modern moral ontology—which give rise to a licit way of discriminating. That kind of ontology is, in fact, "the reality concerned" as Taylor clearly states while explaining the functioning of the strong evaluations.

Being the articulated identity of the agent emerged in relation to that learning process that history is for Taylor, both moral intuitions and the strong evaluations have an authoritative root—something that has received the validation through the historical process and that is constantly open to re-evaluation of single agents.

It is for this grounding of practical argumentation in the wider understanding of moral ontology that an "extension of practical capacities" can be seen as an epistemic gain, that is to say, an increasing of, or a transformation in, the self-understanding which is justifiable, at least potentially, both trans-historically and trans-culturally.

Radical Self-Critique and the Limits of Interpretative Knowledge

The argumentation *ad hominem* based on transitions, suggested as way of implementing a kind of radical critique rooted in self-understanding, is able, in Taylor's view, to support itself because of (a) the indispensable formal orientation to the good, and (b) the historically situated interplay between the moral ontology and the agent's articulation of her hyper-goods.

I argue that Taylor's critical project, that is to say the theoretical edifice he built on (a) and (b), is not able to support radical critique. Yet, Taylor's critique works, to the extent that I think it actually works, as a form of mere interpretative critique. Where by mere interpretative critique I mean a form of practical argumentation ingrained into a meaningful and authoritative horizon of significance which has been inherited, but which cannot itself be subjected to any radical questioning. For in the moment in which practical argumentation, as construed by Taylor, tries to be radical, that is to say, to call into question the authoritativeness of this horizon, it is undermined either by a naïve conception of history as self-conditioned process of the actualization of the potentialities of reason, or by the impossibility of accounting for the continuity and homogeneity of the process of epistemic growth which grounds this form of radical critique.

The study of the moral ontology, as a necessary complement to the trans-historical elements of self-awareness, is meant to be an historical exploration of the "picture which underlies our moral intuitions". Such a picture, according to Taylor, functions as horizon of significance, which, even when tacit because non-articulated, is partly constitutive of our identity. Hence, the work Taylor has undertaken in his analysis of the modern moral ontology can be also summed up as "exploring the frameworks which articulate our sense of orientation in the space of questions about the good". In addition, for Taylor, "these qualitative distinctions, which define the frameworks" are seen "first as background assumptions to our moral reactions and judgments, then as contexts which give these reactions their sense".[97] This means that the horizon of meaning that Taylor articulates and reconstructs is both a conceptual map of modernity, and an authoritative framework. And it is through the authoritativeness of the framework that Taylor intends to secure the

[97] SOS, p. 41.

ground of his understanding of practical argumentation and radical self-interpretation. The success of this theoretical move relies upon the allegedly justifiable vision of the modern framework as the result of (also) "error-reducing" transitions. Modernity, in other words, in Taylor's understanding, can afford a substantive form of radical critique only if its history can be legitimately read as the story of "an epistemic gain".[98]

Taylor's claim concerning the understanding we have of our history as also a process of growth of knowledge we have of ourselves is backed by his focusing on the agent's ability to reflexively and critically engage with her hyper-goods. The process that makes possible reflexive and consequently critical engagement starts, in Taylor's understanding, from the conditions of possibility of judgments of value. A reflective practical judgment implies the presence of first-order desires, and desires of a higher-order. The articulation of the higher motivations and desires—the ones which are indispensable to the agent's identity—places those kind of higher motivations in the foreground. While the mere articulation of the hyper-goods is a critical process in itself because its mere occurrence might bring the agent to reflect on anomalies and mistakes concerning her own (mis)identification with them, when it comes to the judgment of a hyper-good in itself, and therefore to radical critique, the core of the evaluative process becomes the transition. That is to say, a radically critical evaluation starts by articulating the hyper-goods to subsequently interrogate their validity in order to establish "not that some position is correct absolutely, but rather that some position is superior to some other". If such a comparison is able to show that the transition from claim A to B is "well founded" then, according to Taylor, that transition "constitutes a gain epistemically" speaking.[99]

[98] Ibid., p. 72.
[99] Ibid.

The account of 'well founded', according to Taylor, can be provided in the following manner. It is possible, by comparing the views A and B—without using any problematic procedural criteria of validity—to show that A is an improvement upon B, for A is able to solve tensions and problems present in B. Moreover, A is able to show the reasons why B faces those problems; reasons that even people still relying on the competing view are able to see and accept. The account of the passage that leads from B to A constitutes, in fact, a compelling narration, while the same could not be said about the potential opposite narration. "The nerve of the rational proof" then is not as ambitious as to try to reach a universally valid piece of knowledge, it rather "consists in showing that this transition is an error-reducing one".[100]

In the practical field, the spurious concatenation of those transitions constitutes, not history understood as a processes of "diachronic causation"—albeit this kind of history is relevant—rather, it constitutes the emerging of our modern higher motivations. Thus, the second understanding concerns interpretation: not the account of the history of what has happened, but a reading of "what can be called the 'idées-forces' it contains".[101]

Said in a very synthetic manner, what Taylor is suggesting is a form of radical critique that is based on self-interpretation, and yet capable of error-reducing transitions at both a trans-historical level and a trans-cultural one.

Now, when one engages with the details of the examples Taylor provides in order to illustrate the functioning of the transitions, substantial problems emerge.

Taylor provides three paradigmatic examples of transition in order to show that, given two conflicting practical argumentations, or even two whole visions of the good, it is

[100] Ibid., pp. 72-73.
[101] Ibid., p. 202.

possible to recognize one of the two as superior and realizing a gain of an epistemic nature. This is supposed to be the case when:

> The pattern of anomalies and contradiction only comes clear, and stands out as such, from the new position (case 1); the full significance of a hitherto marginalized form of understanding only becomes evident when the new position develops it (case 2); that my present stance reposes on contradiction, confusion, or screening out the relevant only emerges as I make the transition—indeed, in this case, making the transition is just coming to recognize this error (case 3).[102]

From the full descriptions of these three examples, it appears that the only one that, according to Taylor himself, does not rely on something like a criterion of validity is the third one. Yet in the third case he does not address transitions that call into question the authoritativeness of the hyper-goods. In that case Taylor simply explains that when a new option appears on the agent's horizon—new only from the perspective of the single agent—it can lead to the understanding that her self-interpretation was actually given by some fundamental misunderstanding concerning her real motivations. Nothing in that example refers to transitions which are problematic with respect to the validity of the hyper-goods; therefore it is not a good candidate to try to understand whether the *ad hominem* argumentation based on transitions can function as radical form of practical reasoning and critique.

[102] Taylor, "Explanation and Practical Reason" in *Philosophical Arguments*, Cambridge, Harvard University Press, 1995, p. 55.

Concerning the first example, it too does not call into question the validity of the hyper-goods, rather it aims at spotting contradictions and anomalies that undermine our reasoning without providing any reflection on the very sources of the reasoning itself.

Thus we cannot judge the radical nature of Taylor's critique either with the third example or with the first one. For both are inapt to deal with a possible scenario in which what is at stake are the hyper-goods themselves. I am, for this reason, forced to use exclusively the second transition as paradigmatic example for cases in which what is at stake are the hyper-goods themselves.

In the relevant scenario depicted by Taylor, even the parameters according to which something would be considered an anomaly, or a mistake, differ. In this case, A and B face each other from within pre-understandings that are different and that therefore require different and non commensurable explanations. Because the differences pertain first of all their pre-understandings, the explanations A and B, before being confronted with the way they actually deal with the phenomena considered would be completely equivalent, as Taylor himself notices.[103] However, once we look at the way A and B engage with the reality concerned, according to Taylor, we can discriminate between a better understanding and a less good understanding.

If A is able to make sense of what happens in B, while the same cannot be said of B then, according to Taylor, we can consider the expansions provided by A as an improvement for it is able to support an enrichment in our way of making sense of ourselves, and our relation to the world. For instance, from the point of view of a pre-modern understanding of the universe, the "manipulative understanding" of the world would have been impossible to conceive. In fact, its very affirmation would have

[103] Ibid., p. 47.

constituted a contradiction at the core of an order that explains everything in relation to the meaningful place everything and everyone has. On the contrary, in a world where natural sciences are dominant, a vision of reality in which one “understands things in terms of their place in a meaningful order”, is still an available option that does not endanger our self-understanding.[104]

Taylor does not claim that technical control was non-existent before modern science. He simply emphasizes how such pieces of knowledge were marginal, never amounted to an essential part of the pre-modern self-understanding, and that a general account of the universe in manipulative terms would have endangered the pre-modern order.

In an analogues way, our modern aversion to physical pain would be not extraneous to an individual of the medieval period, but, according to Taylor, it would be almost insignificant when compared to the order of the universe and the necessity to respect and reaffirm it even through the horror of torture. The decline of the idea of a meaningful cosmological order, that is to say, of a vision of the good that could be seen as superior to human dignity, contributed to the emergence of a new vision of the good. It was for Taylor “*one* consideration which lends a rational grounding to modern humanitarianism”.[105]

The increased attention to the avoidance of suffering became, as in the case of the manipulative understanding of the inner and outer nature, part of an entirely different self-understanding, an understanding that “draws us because it articulates in a striking and far-reaching form what we already acknowledge in that vague term "human importance"”.[106]

[104] Ibid., p. 45.
[105] Ibid., p. 50.
[106] Ibid., p. 57.

Therefore, even though the new conception of the good is a radical departure from the previous one, "there was *something* which the losing theory had to recognize outside the scope of its original standards".[107]

The recognition of such a something is enabled, according to Taylor, by a "human constant". The constant "consists in our ability to make our way about and effect our purposes" in a given domain, and in the related ability to make our orientation, what I have referred to as a pre-understanding, at least partly, explicit. The articulation of such pre-understanding gives the agent the possibility of judging what set of knowledge is "extending our grasp of the connections which underlie our ability to deal with the world as we do". The prevailing knowledge then has specific characteristics deeply connected to the human constant, and that consist in the ability to orient oneself in relation to the world and to oneself:

> Knowledge of this kind is intrinsically linked with increased ability to effect our purposes, with the acquisition of potential recipes for more effective practice. In some cases, it is virtually impossible to extend such knowledge without making new recipes available; and an extension of our practical capacities is therefore a reliable criterion of increasing knowledge. [108]

Comparing two radically different pre-understandings then relies on an ability that crosses both, that is, the ability of agents of all eras to orient themselves and have some influence on the purposes of that orientation. The implementation of this constant ability though, if my understanding of Taylor's theoretical edifice is correct, is (a) partly dependent on the general horizon of significance for, as seen in section one, self-awareness is always the

[107] Ibid., p. 48.
[108] Ibid.

awareness of an historically situated self, (b) includes the possibility, or even the necessity, to provide "new recipes" at times, and (c) is – "because of these links between understanding and practical ability" – a reliable instrument to evaluate which one is a better form of self-interpretation.

One might think that the fact that this constant ability also relies upon a historically emerged sources gives us a reason to think that the entire theoretical edifice of Taylor is going to collapse. I consider MacIntyre to express this kind of concern when, analyzing the accounts of the emerging of the modern self and its moral sources, he points out that:

> Taylor's […] philosophical account of the modern self and the continuities of its development is presented as if deriving support from those narratives. But those narratives are interpretative, and the interpretation which informs them seems to be itself derived from that same philosophical account for which the narratives are intended to provide support.[109]

Yet, it is simply not accurate to state that Taylor's philosophical account of the modern self-aware agent is founded in—"as if deriving support from"—historical accounts. The foundational element, that is to say, the orientating function of the good is formally justified as I show in section one, and so it is the agent's necessarily engaging in practices of self-understanding. It is nonetheless true that, in Taylor's understanding, the peculiarly modern possibility of radical reflectivity is both linked to, and justified through, historical accounts. But exactly because of the formal orientation, Taylor is legitimated when he writes that he has not made the mistake of invalidating his reasoning by grounding an

[109] A. MacIntyre, "Critical Remarks on The Sources of the Self by Charles Taylor," in *Philosophy and Phenomenological Research*, 54, 1994, p. 189.

allegedly universal truth on an historical limited period.[110] Rather, his account of modernity qualifies the actualization of the formal orientation to the good in a way that is necessary to Taylor's understanding of moderns as self-interpreting beings, but not in a foundational manner. The specific historical actualization is, in fact, inescapable to the "ability to make our way about and effect our purposes", but it is not the fundamental existential condition of the ability itself. In yet other words, even if indispensable, the identity of a given agent, as well as her self-interpretations, vary historically, but what cannot be subject to any variation or interpretation is the necessity to have an identity—this is the formal trans-historical side of the orientation to the good. In this manner, the historical account provides indispensable substantive sources (radical reflexivity, for instance) but those sources are not confused by Taylor with a trans-historical justification for self-interpretation, which is indeed a "constant".

However, the fact that self-interpretation does not undermine itself, or at least has potentially the sources to not undermine itself by falling into a hermeneutic circle, fed by the foundational use of history that MacIntyre ascribes to Taylor, does not mean that self-interpretation amounts to radical critique. In this regard, self-interpretation appears problematic because it is inapt to meet the full extent of Taylor's critical ambitions and his attempt to continue the project of modernity. This becomes particularly evident when we consider how he presents what he defines as "new recipes", that is to say the emergence of notions that once entered into the horizon of significance of the agent, have the chance to pass also the threshold of what is considered authoritative.

[110] "Mais on ne peut pas objecter que cette spécificité historique rend l'argument invalide, comme si je commettais l'erreur de fonder une vérité supposément pérenne sur les conditions d'une époque historique limitée" Taylor writes. See Taylor, "Le Fondamental dans l'Histoire", in *Charles Taylor Et l'interprétation De l'identité Moderne*, edited by Guy Laforest and Philippe de Lara, Paris, Sainte Foy, Les Presses de l'Université Laval, 1998, p. 45.

When self-interpretation is merely seen as an activity aiming at illuminating the sources of our moral practice, it, thanks to the architecture of the strong evaluations, actually works. The chain of theoretical steps, which Taylor has construed, proceeds, at least up to a certain point, unencumbered: a) no legitimated evaluative judgment can be formulated completely transcending from a compelling vision of the good; b) the recognition of this relation to the good does not necessarily condemn practical reason to relativism, but it is the first step toward a better comprehension of the manner we do evaluate; and c) once some clarity has been reached regarding the functioning of the evaluation and the deep motivations associated with it, such motivations can be rationally engaged in virtue of a form of local reasoning based on transitions which is the final capitalization of the interplay between the trans-historical and the historical.

Yet such interplay between the authoritativeness of the historical sources and the formal elements of the self-understanding are unable to account for the passage from what is meaningful to what is authoritative, and, at same time, comply with Taylor's declared goals: keeping the judgment of value within the realm of what is epistemologically valid while leaving space to the agent to call into question even her higher motivations.

In the example of the second kind of transition, that is paradigmatic of a radical shift in the way the agent understands her form of life, what changes is the authoritativeness of the framework. Consequently, Taylor's focus is on the ways in which, and the reasons why, this framework changes.

When we consider more closely the example he provides, we see that, in spite, of the radicality of the shift described, it is not an example concerning the emergence of one or some higher ethical standard. Instead, it is an illustration of how marginalized notions, that is to say, something already meaningful, once recognized as capable of providing a

better sense of our form of life, are able to bring about a revolution concerning what we understand as already worthy . There is no point in Taylor's narration in which he describes the historical emergence of a single hyper-good. Even equality is, in his narration of modern history, the slow evolution of pre-existent intuitions around the conception of oneself as something worthy.

This absence of a new hyper-good in the practical field could gives rise to the suspicion that, according to Taylor, human history came about with all the hyper-goods already present and that the adoption of new or marginalized notions he refers to is the manner through which we improve and refine our understanding of those hyper-goods—a sort of gradual comprehension and adaptation to the full potentiality of something that, at least in an embryonic stage, is already there.

Within this vision, we moderns have the chance to engage those hyper-goods in a meta-theoretical manner and therefore have the possibility to bring about a further improvement concerning their actualization; still we cannot forge or actually abandon any of them. And by forging I do not mean the picking *ex nihilo* from the normative horizon—an option that Taylor shows to be inconsistent. On the contrary, I simply have in mind a historical process guided by real possibilities for actual alternatives to emerge, rather than the unfolding of a process of gradual illuminations of principles already existing in some form.

The fact that the hyper-goods appear to be all already there albeit inchoate translates into the understanding of modernity not as a new framework shaped also by a new understanding of the good, but merely as the result of the collective work to articulate, as clearly as possible, something that have always been present in history, but

not (yet) in a clear form. In this scenario, history, at least with respect to the good, takes the shape of the only real agency involved.

Such an understanding of history does not entail a real possibility of radical critique. For in order to engage moral sources in a radically critical manner, one should be history itself, as it were. Furthermore, it also cannot make sense of the real plurality of cultures as all the different conceptions of the good would be intermediate stages on the way to approximate the full unfolding of the same range of goods in history. Taylor's critique, in this scenario, would be a simple interpretative critique, that is to say, the gradual reflexive self-comprehension of the process of flourishing of the sole form of life that has ever existed. The passages from A to B and then from B to C would then be justified as error reducing, but only given the assumption that history has originally provided us with the right hyper-goods.

History, thus, more than the ground for an uncertain growth of self-understanding, would look like a metaphysical assumption. An assumption that is probably unavoidable for an argument to the effect that merely interpretative critique suffices. But such an assumption is exactly the kind of thing that needs to be called into question when it comes to radical critique.

This, though, is not the only possible reading of Taylor's work. In fact, it is probably a very uncharitable one, especially if one recalls the theoretical efforts Taylor has made in *The Sources of the Self* to provide an account of modernity as something which has developed in continuity with the past but that also presents us with unprecedented elements. In the light of this, one could defend Taylor against my critique by claiming that it is based on a wrong interpretation, for he would never embrace such a naïve understanding of history. By implication, there must be actual space for "new recipes" even

in the practical field. It must be the case that the affirmation of new notions, as, for instance, the one of inwardness, can alter so much the comprehension of what we already see as worth that in the end "the reality concerned" (the hyper-goods that constitute our moral ontology) is transformed too.

If this is the case, then, there is space for real transformations. And if there is space for real transformations in Taylor's understanding of history, then the new fragments of the meaningful that contribute to those transformations can reach the domain of authoritativeness only in two ways. They need to be either measured against the pre-existing authoritativeness (the old hyper-goods) or to pass the test of "extending our grasp of the connections which underlie our ability to deal with the world as we do".

If we are in front of a new notion it is unlikely that the agent is going to find such a notion compatible with the old hyper-goods, and more generally, with the whole framework the agent is in the position to adopt in order to produce a legitimated judgment. For, if this were the case, the notion in question would be a minor transformation, a mere variation on an already valid theme. The likely scenario—the one in which the old hyper-goods are of no use—is the one I am interested into.

In this scenario the old set of hyper-goods and the whole framework they are part of would be simply of no help, and it would be necessary to appeal to the constant ability "to make our way about and effect our purposes" even beyond the standards embedded in one's own pre-understanding. Yet, this amounts to a switch from the authoritativeness of historical sources to a constant ability, which does not really work as valid epistemological guidance—and this is the very point I want to make. When the agent ventures into the field of what starts to be meaningful but has not yet reached the status of authoritative, the necessity (dictated by the contingent circumstances) to recognize something as a

potentially viable option—even if such an option does not comply with the original standards—must mean that some judgments are produced drawing on notions which are not (yet) legitimated. In spite of the absence of authoritativeness, such notions must be capable to tentatively drive the reasoning in question and actions of practical relevance, for on the moving front of history we do not have ready-made valid alternative options waiting for us once the old ones fail.

Taylor tries to solve the difficulty of bridging the transition from one authoritative framework to another by invoking a constant ability that is allegedly capable of keeping the transition itself under the umbrella of what is epistemically justified. Yet, he never really explains why the ability "to make our way about and effect our purposes" has by itself an epistemic authority rather than, for instance, being merely capable of bringing about a change in meaning. Taylor's inability to give an account of the reasons why the knowledge produced by the mentioned ability is immediately authoritative has as a consequence the fact that the very transition from a pre-understanding A to B cannot be seen as also epistemically justified.

Moreover, the structural difficulties Taylor's theory faces in attempting to justify the way in which new notions become authoritative has another corollary: when describing the transition from vision of the good A to vision of the good B as an epistemically justified transition, Taylor ends up hiding the actual historical transition more than disclosing it.

When comparing two pre-understandings, A and B, Taylor construes the study of the transition as a retrospective study. In his example, A and B are already existing understandings, and each depends on a different authoritative vision of the good. At the same time, the account of the emerging modern sources Taylor describes, which covered a period of two millennia, unambiguously suggests that the historical passage from one pre-

understanding to another is a piecemeal process. In addition to the retrospective form of argumentation, and the piecemeal proceeding of history, it must be also remembered that according to Taylor himself meaning and norms "are not just in the minds of the actors but are out there in the practices themselves, practices which cannot be conceived as a set of individual actions, but which are essentially modes of social relation, of mutual action".[111] Keeping in mind these three different elements, it is very difficult to see how the description of the transition Taylor provides can be a reliable description of how historical transformations occur at the level of the authoritativeness of the framework.

Consider the following example. As agent who is not yet used to egalitarian notions, I might end up in a circumstance in which equality, or something that does not even have that name yet, enters my horizon of significance. Perhaps in order to sustain myself I am forced to work in a factory in which, regardless of gender, we are all equally exploited. The way I come into contact with this notion, then, is not an intellectual engagement, but has to do with my daily practices. Probably, if questioned, I would claim that as a woman I am less capable than a man, even though in the meanwhile my practice would suggest otherwise. Once my experience has been informed by what I do in addition to what I think I should do, probably my understanding of myself will start to be problematic. Given the contrast between my compelling hyper-goods and my actual practices, it is likely that I will experience such a contrast as something painful. I would probably think of myself as a person not really capable of living in accordance with what I always thought a woman should be. Yet equality at that point is part of my aware experience. This notion, even if in a deeply problematic manner, starts to make sense to me.

[111] Taylor, "Interpretation and the Sciences of Man", in *Philosophical Papers*, Vol 2. Cambridge, Cambridge University Press. p.36.

At a certain point, this hypothetical agent herself, or more likely an external observer, engaging with the new social practices and meaningful notions related to them, can even recognize the transition from inequality to equality as an epistemic gain. Once equality and inequality are there as fully formed notions, one can, in fact, compare, and even point out good reasons as to why the transition is an epistemic gain. Yet this retrospective account of the transition will say nothing about (a) how that notion of equality entered first the horizon of significance and then reached the status of authoritativeness, and (b) how self-understanding changed at a personal level.

The fact that the kind of account Taylor presents is not able to provide an explanation concerning point (a) and (b) is a problem because in order for his constant ability to work as he intends it to work—that is by making our way about while ensuring that that very way is epistemically justified—he must be able to also explain both how significant notions reach the status of authoritativeness and how this transformation is either brought about or mirrored by the agent's self-understanding—exactly point (a) and (b). Those two points are, in fact, nothing more than questions about the functioning of the constant Taylor refers to. For it is that constant ability that is supposed to work in such a manner to produce a justification for (a) and (b).

Insomuch as Taylor is not able to explain (a) and (b) as processes which are epistemically justified, then he cannot account for the homogeneity of the entire process he defines as epistemic gain. That is to say there are gaps in the authoritativeness of the process of our self-understanding's growth. Our history, even when it is not considered as causal history, but only as the history of the *idée-force* which shape our self-understanding is still not a homogeneous process of epistemic growth. Consequently, a strategy which is

grounded in the transitions Taylor suggests would be insufficient to justify the kind of epistemic growth that is required by the form of radical critique analyzed here.[112]

And Taylor is indeed not able to explain the processes (a) and (b). The passage from the mere significance of a notion to its authoritativeness—given the theory of the strong evaluations Taylor works with—can be produced either as an initial experimental transgression which, in virtue of social practices, will eventually be organized in the coming paradigm, or as conversion.

In the first scenario, our ability to make our way about is not only not justified by an authoritative framework, but it works—in the measure in which Taylor's understanding of the strong evaluations can possibly allow it to work—because it can experiment also while not justified.

A post-facto account of the transition, in this kind of scenario, could be described as an epistemic gain only on the ground of the new authoritative framework. However, this cannot count as a sufficient explanation as it is Taylor who invites us to focus on the transition *between* authoritative frameworks and not *from* the perspective of one authoritative one—otherwise he would not need to recur to the notion of a constant ability in the first place.

Concerning the second scenario, the one which would resemble a conversion, it too, once the conversion has happened could provide us with sources to see the very conversion as an improvement of our epistemic plight, but it too, would leave open the question about the reasons why we came to see the new authoritative framework as superior.

[112] As mentioned in footnote 3 of this chapter, the main concern of this part of my work is to understand whether a radical form of critique can be sustained in the manner presented here. The fact that the forms of transition taken in account here is not capable of accounting for the kind of continuous epistemic growth which would be capable to ground a form of practical reasoning which draws also on substantive sources is, therefore, relevant beyond the issue of what Taylor's own position ultimately is.

The transitions pertaining the shifts in our understanding of our form of life, therefore, do not happen on an epistemically legitimated ground, they are, instead, even obscured by a reasoning that either focuses on elements which have both already the status of competing views, or by looking back in a way that takes the outlook that superseded a previous one implicitly as authoritative.

If my analysis according to which Taylor has no functioning instruments to deal with point (a) and (b) is true, then the history of the emergence of the modern self and its moral sources cannot be described at the same time as the history of also a continuous epistemic gain. As a consequence, radical critique is structurally impossible within Taylor's theoretical edifice. Its possibility, in fact, relies exactly upon the vision of the modern framework as the result of (also) "error-reducing" transitions. The absence of an authoritative framework able to evolve while keeping its validity at a trans-historical level entrusts self-interpretation and its critical function to a vision of the good that cannot be radically questioned.

While self-interpretation, as construed by Taylor, fails as ground of radical critique, it can still work as instrument aimed at illuminating our modern moral sources and the way we use them to formulate judgment of value.

However, Taylor's work, even when seen as a mere form of interpretative critique faces relevant issues. Issues that are not structural as the ones I have just pointed out—at least I do not read them as such—but that still constitute a problem.

Nussbaum's criticism, for instance, starts out from Taylor's account of modern identity in order to subsequently arrive at the possibility to use his historical account as a justification for practical rationality. She notices that the reconstruction that Taylor

provides is not accurate. “Taylor has not told the entire story” she states.[113] In spite of his clear attempt to do justice to the complexity of the modern western identity, Taylor, according to Nussbaum, has left out of the picture many cultural aspects that are not originally part of the western tradition, but that, nonetheless, shaped the sense we have of ourselves. And the same can be said, in her view, of all those ideas that have not found any positive resonance in modernity.

In particular, Nussbaum points out that ““our” history had better involved, for the purposes of moral and political reflection, some understanding of African, Asian, and Latino traditions”. In support of her view she notices that the influence of those traditions, even if “less easily accessible to a historian with Taylor’s text-centered methods” might have had a big influence on the conceptions he analyzes. Moreover, a modern society is exposed not simply to domestic politics, and this has an influence on our ethical life.

In spite of all these problematic elements, Nussbaum thinks that the fact that Taylor does not write the whole story, in itself, does not undermine his own goal since his is an account of the way we understand ourselves. The element actually problematic concerns the way Taylor uses his historical account to justify moral judgment.

For instance, Taylor has provided an account according to which the passage from the ancient Greek aristocratic understanding of political life to a view of human dignity as universally shared is a gain generated by tensions already present within the ancient conceptions themselves. Still, according to Nussbaum, if this very passage is explored further, it proves to be not so straightforward. According to her, there are different issues. First of all, it is important to distinguish the “real and urgent difference between progress in understanding and the imposition of political or cultural power”. This issue within Taylor’s

[113] M. Nussbaum, “Our Pasts, *Ourselves*”, in *Philosophical Interventions: Reviews 1986-2011,* Oxford University Press, New York*,* 2012, p. 89.

account of the passage does not find an obvious answer. Focusing on the way Taylor deals with historical transitions, she writes:

> Let A be the Aristotelian view that […] the human being combines capability with limitedness and vulnerability. Let B be the Augustinian view that our fundamental situation is one of sinfulness. To me, it seems most illuminating to narrate the story going not from A to B, but from B to A: people at some time come to realize that their basic situation is not original culpability, but capable finitude, and this is a gain in understanding.[114]

Yet history has gone, as Taylor suggests, towards the Augustinian view. This very historical passage together with Taylor's idea that a better self-understanding can be a reliable meter to measure possible gain, fortifies her suspicion that the power of religious institutions has played a bigger role compared to the "sense that the Aristotelian view cannot explain all the ways in which life goes badly".

Even while pointing out this problematic issue about power, Nussbaum adds that the argument can be adjudicated in different ways and that her observation, by itself, does not undermine Taylor's argument about transition, "the enterprise of justification does not need to collapse". Yet she does not "see exactly which way Taylor favors".[115]

While Nussbaum reckons Taylor's work as lacking sources capable to address problems associated with power, I think that Taylor's work, while actually marginalizing issues of power, addresses a very similar concern when he deals with physical punishments. He, in that occasion, recognizes that "[t]here are other, independent grounds in modern culture which have made us more reluctant to inflict pain. Some of them may have sinister

[114] M. Nussbaum, "Our Pasts, *Ourselves",* in *Philosophical Interventions: Reviews 1986-2011,* Oxford University Press, New York*,* 2012, p. 91.
[115] Ibid.

aspects", yet, he goes on, we must recognize that among those independent grounds there are also considerations of moral character about motivations, that is to say, the conditions of intentionality. Once those motivations have been articulated, their explanatory force can, even if not always, be validated or invalidated against different ones.[116]

While this kind of answer would be not sufficient to justify a form of radical critique, I read it as fully compatible with a form of interpretative critique that has renounced to be radical, but it is still capable to address Nussbaum's concerns. This, at least, if we consider those concerns as not extending till the point of questioning the very structure of self-interpretation.

Skinner's reading of Taylor's work is a sharp collection of criticisms which all derive from his own historical reconstruction of modernity. According to Skinner, the values Taylor reconstructs and presents as constitutive of the modern identity are not as unproblematic as he claims them to be. In Skinner's reading, the selected account of modernity Taylor provides appears more a vindication of our bourgeois lives rather than an accurate account of the modern conception of the good. For instance, when it comes to the description of the affirmation of the ordinary life as something that orbits essentially around family life, Taylor decides to simply ignore the doubts that have been casted on the place of the family in liberal societies.

Skinner notices that Taylor fails to take in account that "[a]lthough the ideology of modern liberalism has always proclaimed the value of foredoom and equality, modern liberal societies have always been underpinned by a sexual contract in which these ideals have been systematically violated".[117] Ignoring this kind of criticisms of the values he takes

[116] Taylor, "Explanation and Practical Reason" in *Philosophical Arguments*, Cambridge, Harvard University Press, 1995, p.50.
[117] Q. Skinner, "Who Are "We"? Ambiguities of the Modern Self", in *Inquiry*, v. 34, (2), 1991, p. 141

to be pacifically accepted and recognized, Taylor has been, in Skinner's view, dismissive of "Marx's insight that a major source of alienation from our true identities may lie in the character of our daily commitments themselves".[118] In addition, closely analyzing the historical account Taylor's provides, Skinner notices that "the more we acquaint ourselves with the kind of causal story that historians like to tell—but Taylor elects to ignore—the more does such an affirmative stance seem impossible to uphold like to tell". Looking into that kind of stories, according to Skinner, one would be surprised about the number of groups of power that "had interest in patronizing and encouraging the growth" of the value concerning family life.[119] Since Taylor presents himself as providing an account which wants to be seen also as an epistemic gain then, even while not considering his account as causal history, it is nonetheless relevant to wonder about "why this type of causal story tends to subvert Taylor's philosophical argument".[120]

Skinner's criticisms are on target when it comes to pointing out the surprising agility with which Taylor avoids to pay attention to the role power, interest, and ideology play in history, but he, in turn, seems to ignore the actual philosophical point at stake here. When Skinner points to criticisms concerning family life and the possibility for it to be an oppressive institution rather than a way to affirm freedom and equality, he is indirectly proving Taylor's point about modernity. Skinner is both holding that freedom and equality are indeed seen as worth even by theorists who do not share his same understanding of the institution of marriage, and reinforcing the modern practice *par excellence*: reflective self-understanding. It is to this very activity that Taylor entrusts the possibility to fruitfully deal even with the issue of power and ideology. His historical narration aims at proving that

[118] Ibid., p. 142.
[119] Ibid., p. 144.
[120] Ibid., p. 145.

we moderns are those kinds of beings who, again and again, can interrogate themselves about how fully we comprehend the hyper-goods that we allow to drive our judgments, and how coherently we are applying them.

Yet, as in the case of Nussbaum, when we move from the perspective of simple interpretative critique to radical critique then, power relations and an accurate account of history assume a different value. That is to say, their study brings to light what, from the perspective of an already functioning horizon of normativity, is no longer visible.

While Taylor's work is able to address some of the structural problems present in Habermas' discourse ethics, it is, in turn, unable to provide feasible answers to the concerns which move both researchers: how to provide a valid foothold for radical critique and, by implication, for the project of modernity.

Since both Taylor's and Habermas' work are paradigmatic of the two main ways in which this issue has been faced, it would be legitimate to think that it might not be possible to ground radical critique and the project of modernity.

In some ways, this suggests that we should abandon the modern ambition to answer to the question of modernity.

Yet, in order to save the spirit of modernity, perhaps what must be questioned is the question of modernity itself, or how the project of modernity has itself been interpreted.

Foucault and the Re-conceptualization of the Project of Modernity

The emancipatory project of modernity has been addressed essentially in terms of knowledge. Radical critique, thought of as the tool to implement it, aims at providing modern reason with a reliable manner with which to mediate its self-comprehension. The knowledge produced through this process is needed, not only to satisfy questions about limits and enabling conditions of truth, but it is also, and above all, supposed to serve as a guide within the moral field.

Knowledge and freedom find themselves intertwined in the question of how modern reason justifies the aims it legislates for itself. This means that the decisions about how to live and how to shape public life are seen as necessarily linked with the (either formal or substantial) knowledge that is meant to provide a justification for those decisions—the aims reason legislates for itself.

The necessity to look for a foothold capable of grounding that knowledge is an essential assumption of the paradigmatic works I have analyzed in the previous two chapters. We have seen that these works do not question this assumption itself, however, the difficulties associated with this task have been the object of attentive studies. Habermas and Taylor themselves are well aware of the problems that the identity between reason *qua* foundation and reason *qua* object to be investigated are causing. The doubling of modern reason creates, in fact, difficulties concerning the process of validation. More importantly, such a doubling raises the question about how to investigate, within the empirical, the conditions of possibility for such validation without ending up invalidating the investigation itself. For, given our finitude, the overlapping of empirical conditions and transcendental ones would undermine any radically critical project. The knowledge

produced in this manner, in fact, is the one of a divine mind which is transparent to itself. Consequently, Taylor and Habermas have both tried to look for a foothold that is partly external, as it were, to reason itself, so that reason could have the needed mediation in order to investigate itself.

Habermas has individuated the sources for this task in a communicative model of reason. In this model, the reason of the single knowing subject is decentered in a formal system of norms. Those norms are, at the same time, organic to the functioning of reason and made by the community of speakers. They are both reason itself and something other than it. As a result, the actual speaker can recognize 'out there' those norms that are required to establish, in Habermas understanding, the validating procedures for reason to follow. Reason in this manner, allegedly grasps a part of itself in something that is not exactly itself but is not extraneous to it either. This, from Habermas' point of view, is a sufficiently post-metaphysical foothold for reason to deliberate in a valid fashion about its own aims.

Taylor, instead, thinks that even if the discourse ethics which derives from Habermas' communicative model of reason can be a valuable instrument to apply to the practical field, it cannot make itself autonomous from the idea of the good already embraced by the speakers. Hence, Habermas, from Taylor's point of view, is simply not able to do what he claims to do. On the contrary, discourse ethics, if anything, is a specific product of the moral sources already valued by moderns.

The misrecognition of the fact that the articulation of whatever kind of judgment has an orientation, which Taylor defines, very broadly, as moral, hinders the understanding, and the consequent assessment, of said orientation.

From his perspective then, the moral orientation and the related hyper-goods that are already in place – both in the social fabric and at the level of the personal identity – are the sources which must be looked at in order to shape the needed foothold. The fact that reason cannot reach any direct and full knowledge of itself does not mean, in Taylor's understanding, that reason, through the account of its historical evolutions (learning processes), cannot achieve the knowledge of what it came to be for us moderns: moral ontology.

The investigation of both who we are for ourselves, and the composite processes which brought us to be who we think we are, according to Taylor, brings to light a kind of interpretative knowledge which is supposedly able to justify the aims reason has set for itself along its historical substantiations. Reason knows itself, as it were, through the historical and social articulations which the moral agents have selected and tested. That is to say reason finds a way to mediate its own knowledge through articulations that, as a result of being made and then selected by the human consortium, turn out to be also valid.

With the previous two chapters I have shown how these two promising and opposing routes of providing an answer to the question of modernity have failed in regard to their own aims, due to architectural problems of these approaches. In this chapter I look at a re-conceptualization of modernity which both highlights how those attempts were doomed as they tied the fate of the modern project of emancipation to knowledge alone, and seeks a new strategy to enable that very same project.

Modernity, according to Foucault, can be accessed through multiple entrances. When accessed through the one of knowledge and truth, the critical function which that very entrance enables is in the position to raise a certain range of questions but not others. The questions allowed by that perspective will be only about the limits of that knowledge

and the way it is justified. This perspective is the one Habermas and Taylor use, and it is the only one they reckon as practicable or even desirable in order to revitalize the modern emancipatory spirit.

Foucault thinks that very perspective is both ill-suited and counterproductive given the liberating effects that radical critique wants to have. This is because investigating from that point of view, while aiming at being radically critical, they cannot but end up with a version of what Foucault calls the "analytic of finitude". That is to say a mistake that concerns a switch between the empirical and the transcendental which is triggered when a piece of our finitude rises to the rank of knowable condition of possibility—radical critique.

Foucault does not specifically address either Habermas' or Taylor's works, but, as I will show later, he criticizes the archetypical forms of said mistake which can be identified in the logic which sustains both their works, and in general any form of radical critique that finds its *whole* justification in a piece of valid knowledge.

Even if Habermas has performed an important shift from the conscience of the subject and its reason to the rational and decentered structure of language, exactly in order to avoid said mistake, his model of communicative reason still reproduces the substituting of the empirical with the transcendental. The community of speakers needed is the one which supposedly knows itself and the manner in which it can liberate its own communicative competences (the ones that work without objectifying); yet the community actually involved in dialogues is the empirical one which is not in the condition to reach transparent self-knowledge. Therefore Habermas allows himself to construe discourse ethics on a fictitiously transcendental community, as it were. For it is such a community of speakers that serves as a condition of possibility. This means that Habermas works as if he

could switch without any problem from the empirical to the transcendental level, as is presupposed by discourse ethics.[1]

Taylor, instead, employs a specific idea of a unified history which supposedly is the discovered origin of both modernity and the modern subject, but which is actually a presupposed condition of possibility of the interpretative knowledge he uses as foundations for his own form of radical critique.[2]

In other words, while Habermas takes for granted something about the subject of the investigation, Taylor is guilty of the reverse mistake of taking for granted something about the object of the investigation.

While highlighting the characteristically modern mistake(s) they make, Foucault is in no way suggesting that critique can do without any form of knowledge, or that looking for the limits of knowledge is of no use. What he is suggesting, instead, is that emancipation cannot derive from knowledge alone because knowledge, particularly when valid, is not innocent.[3]

One of Foucault's main discoveries is that knowledge as true discourse—therefore not simply its distortions—*also* has subjugating effects. If it is reason and its forms of rationalizations that generate costs that are too high to bear, then it is from those effects that one must try to defend oneself. And this kind of defense cannot be generated from the

[1] About this very point see also Poster M., *Critical Theory and Poststructuralism*, Ithaca, Cornell University Press, 1989.

[2] About Taylor's understanding of history see R. Fillion, 'Foucault contra Taylor: Whose Sources? Which Self?', in Dialogue (Fall 1995), pp. 663-74; and Skinner, "Who Are "We"? Ambiguities of the Modern Self", in *Inquiry*, v. 34, (2), 1991,

[3] About this point see the following quotation concerning liberating practices: "These struggles are not exactly for or against the "individual"; rather, they are struggles against the "government of individualization." They are an opposition to the effects of power linked with knowledge, competence, and qualification—struggles against the privileges of knowledge. But they are also an opposition against secrecy, deformation, and mystifying representations imposed on people. […] but neither is it a skeptical or relativistic refusal of all verified truth. What is questioned is the way in which knowledge circulates and functions, its relations to power. In short, the regime of knowledge [savoir]." Foucault, *The Essential Foucault,* Paul Rabinow and Niklas Rose ed. New York, New Press, 2003, p.130.

good side of reason, so to speak, neither can one give up on reason altogether. Under such conditions, the question becomes one of choosing a different strategy.

The form of critique Foucault intends to shape and practice "is the movement by which the subject gives himself the right (*le sujet se donne le droit*) to question truth on its effects of power and question power on its discourses of truth".[4] And this questioning does not start as consequence of either a purely intellectual curiosity, or as an "original aspiration". It emerges as a reaction to local dynamics which produce effects that while being experienced as negative cannot be adequately addressed by drawing on the already established legitimated sources.[5]

While one cannot anticipate which aspects individuals will experience as negative or problematic, since these are effects of particular constellations, how to give shape to the decision to question and possibly change the game of the interactions of truth-power is, according to Foucault, exactly the work of critique as the task of modern philosophy:

> But what, then, is philosophy today — the philosophical activity, I mean — if it is not the critical work that the thought brings to bear on itself? And in what does it consist, if not in the endeavor to know how and to what extent it might be possible to think otherwise, instead of legitimating what is already known.[6]

Formally, critique is still what Kant thought it to be: the experience that thought has of itself. Yet as soon as one zooms in, all the terms of the relations appear modified to an extent that reveal a tension with the original understanding of *Kritik*. The experience

[4] Foucault, "What is Critique", in *The Politics of Truth*, The MIT Press, Boston, 1997, p. 47.
[5] Foucault, "What is Critique", in *The Politics of Truth*, The MIT Press, Boston, 1997, p. 74-76.
[6] Foucault, *The History of Sexuality*, Hurley tr., New York, Pantheon Books, Vol. II (The Use of Pleasure), 1984b, pp. 8-9.

Foucault talks about is productive. The aim of Foucault's form of critique is not to secure the limits within which reason can unproblematically push its own knowledge, but it is to contain reason's hegemonic effects. The only way to do that, according to Foucault, is by trying to think otherwise.

It is due to this very aim that we find ourselves being confronted with a new necessary study of the limits of thought. Constraints and conditions of thought are investigated in its institutionalized forms (archeology), and in its effects of power (genealogy) in order to make space for possible alternatives. The dynamics of history, the webs of social relations, and the relations the subject has to itself are closely investigated in search of those limits.

This means that Foucault sets himself an extremely demanding aim: maintaining a formal philosophical orientation which deals with the limits (both constraining and enabling) of truth, while making room for a practice of liberation which is partly enabled by those limits without being justified by them.

This task does not allow for immediate and simple results, and pursuing it has required different and multiple adjustments triggered not only by the problems encountered in the process of exploring, testing, and verifying the conditions and constrains of the possible, but also by the kind of retroactive modifications this very practices have caused. For these reasons, archeological and genealogical studies, that is to say the tools shaped in order to spot, modify, and loosen the disciplinary effects of modern reason, have not traced by themselves a linear and smooth process. In addition, what emerges from exclusively those archaeological and genealogical studies is a work that even while gesturing toward the possibility of thinking otherwise – for they are able to highlight those spaces where knowledge was not yet valid knowledge, and power is seen in its

dynamic relation with freedom – they abstemiously arrest themselves on a sort of threshold.

What those studies entail, in positive terms, has become clearer, perhaps to Foucault himself, only when he, bringing together the several threads of what he had been doing in the previous twenty years, with his work on *Critique et Aufklärung* points to a way to capitalize on his previous studies and go beyond them.

That is the moment when the rough route of Foucault's practice of critique became unambiguously clear, providing the support to read the relentless work of both disassembling erroneous assumptions (as for example the one on which radical critique rests), and pursuing history and social webs as efforts made already in order to get the present to take the step to maturity.

Looking at Foucault's work retrospectively, in fact, it appears that his skepticism, and the freedom to exercise it, even against the very concept of radical critique, have been employed not in order to end up "with the bare abstraction of nothingness or emptiness" that "cannot get any further from there" as the old Hegel would put it, but in order to try to restore the project of modernity from its very roots: *Aufklärung.*

It is in virtue of both the *pars destruens* and the *pars construens* of Foucault's critical practice that in this chapter I present his work as a specific re-conceptualization and substantiation of the emancipatory spirit of modernity. In order to both differentiate between Foucault's critical work, and the ones which draw on a (conventionally) normative justification, and stress the value of Foucault's critique of critique, I will refer to it as hyper-critique.

In a manner that intersects both the constructive and deconstructive sides of his critique, Foucault shows that the formal philosophical work and the decision to be, act, and think differently are depending on each other.

The formal side of his work, by itself, does not legitimate any critical practice that aims at a real change. Instead, that formal side is one part of larger critical practice. But neither is it the case that a mere will "to think otherwise" can, by itself, give shape to a real possibility. Critique is then, for Foucault, the point where the formal philosophical work encounters a self-grounding decision to try to shape the possible—that is a modern ethics based on a transformed study of limits and enabling conditions of thought.

My overarching reading allows me to look at Foucault's whole work identifying what I take to be the different components of his critical practice, and to show their functions as parts of a single long-lasting partisan endeavor—"the endeavor to know how and to what extent it might be possible to think otherwise, instead of legitimating what is already known".[7]

My intent is not to deny the variety of discrete aims, and the, arguably, discrete difficulties Foucault faces with each of his writings and moments of his intellectual production (archeology, genealogy, and care of the self), but to analyze and stress the specific conception of the philosophical practice which, according to my reading, has guided him throughout. In fact, an interpretation aiming to trace a continuity of content and/or method would be easily proven wrong, but if the option of the ethical stance animated and clarified by critique is as reliable as I think it is, then what must be sought out is a single logic, an intent of philosophical character detectable in each of the works, but not exhausted in any of them.

[7] Foucault, "What is Critique", in *The Politics of Truth*, The MIT Press, Boston, 1997, p. 32.

The several elements which compose Foucault's "diagnosis of modernity", or later "ontology of the present", even while (willingly) not constituting a critical *theory*, still has to be considered an encompassing and analytically endowed critical *practice* that embraces the spirit of modernity rather than dismisses it.

This re-conceptualization of the project of modernity that Foucault offers, while not providing an answer to the question of modernity—it turns out the be the wrong question— is able to keep constantly open the possibility to think and act in a modern way; the maturity of which Kant first wrote about.

The modern ethics that Foucault shapes is, in my reading, capable to explain why radical critique fails at keeping up with its own aims for it embodies a characteristic trap of modern thought in which its logic is enclosed: the analytic of finitude.

In addition, the critical practice that animates that ethics encompasses a project of emancipation that, uncoupling itself from the attempts to answer to the question of modernity, links its fate to a modified study of the conditions and constraints of thought. This modified study of these limits, while willingly giving up on the project of making critique a science of the limit, succeeds in turning it into an informed practice of the limit. Such a practice gives a fresh and promising chance to the modern project of emancipation.

Furthermore, as I will show with the last part of this chapter, hyper-critique can defend against the main criticisms that have been levelled against it, and it can be shown that it is not falling prey to the mistakes it intended to avoid.

The Trap of Modernity and the Constant Possibility of its *Ausgang*

According Foucault, a double path crosses modernity: production and discipline, on the one side, and a sort of exceedance, on the other. The processes of a rational domestication of life are at work in the production of the subject, but, at the same time, they are also the only 'places' the autonomy of the subject can find opportunities.[8]

This intrinsic ambivalence opens up a double issue which is played out within the tension between the necessary determining of the subject and its similarly necessary antagonistic attitude.

Consequently, an understanding of critique as a liberating practice can find a supporting point only by turning towards and against that side of modernity which function as determining force. It is the attentive study of the discourses and practices which shape the understanding of ourselves and our conducts that, according to Foucault, shows both the limits of that liberating practice, and, in so doing, also its enabling conditions.

This means that the practice of critique needs to work in relation to its specific targets, for it is only by starting from the limits generated by a specific problematic condition that one can also pinpoint the potential manners to overcome these limitations.

From a philosophical point of view, the main obstacle towards a more fruitful understanding of our time, and consequently towards our possible emancipation, comes from what Foucault indentifies as the "analytic of finitude".[9] As a consequence, in this

[8] Foucault clearly sates: 'A power relationship can only be articulated on the basis of two elements that are indispensable if it is really to be a power relationship: that "the other" (the one over whom power is exercised) is recognized and maintained to the very end as a subject who acts; and that, faced with a relationship of power, a whole field of responses, reactions, results, and possible inventions may open up'. Foucault (1982, p. 220).

[9] It is worth noticing here that the concept of emancipation is not necessarily tied to the one of progress. The processes of liberations are constant and experimental. Given the impossibility for the subject to reach a complete understanding of the passive/active games of subjugation/subjectification of which she is the effect,

section, I provide both a reconstruction of it as a target, and an alternative view that aims to clean the field of investigation from it, in order to subsequently have that field opened up to practices of critique. The following sections, then, provide a reconstruction of the theoretical tools that play together in order to make that practice a feasible and potentially successful strategy.

If modernity is, as said in the introduction to this chapter, a place with several entrances, then, according to Foucault, Kant has directly or indirectly built them all. His comments on Kant's answer to the question concerning *Aufklärung* highlight the presence of, at least, two Kants: on the one hand, the Kant which prepares the terrain of modernity as the reign of an analytic of finitude; and, on the other, the Kant who originates also an antagonistic modernity.[10]

While the latter is completely involved into the question about the *Aufklärung,* which Foucault appropriates as research of a space for emancipation within the present, the first Kant—the Kant of the Kritik that Foucault defines as the father of the analytic of truth—originates a tradition which is, instead, a target.

Modernity, explains Foucault in *The Order of Things,* is in an anthropological sleep. Kant has awakened himself from the dogmatic slumber to consign us to a new prolific illusion that finds its theoretical embryonic start in Kant himself and lives on in the most recent conceptions of radical critique.[11] This illusion is exactly what, at times, is defined as

---

there is no criterion for speaking about freedom as a status or about any improvement. There are just the unforeseeable and in some regards untamable effects of contingency which are capable to modify the economy of the possible within which the subject moves and is.

[10] The idea of two Kants and of two modernities entails also a division between the Enlightenment of the progressive march of reason and the *Aufklärung* of the present as a difference within history. And that is why here the two terms are not used synonymously.

[11] In the last part of this chapter, I will came back on this point clarifying that for Foucault the Kant of the first Kritik still keeps the two elements separated—he actually postulates the separation there—yet later on, in the Logik, by raising the question about what is Man, he ends up overlapping them.

"the anthropological turn" or "analytic of finitude" that is the trap in which critical thought is caught.

The illusion emerges through the doubling of the transcendental subject in its objectified and therefore knowable empirical version. Such doubling, what Foucault respectively defines as subject of knowledge and object of knowledge, inaugurates the birth of Man. As soon as the subject doubles itself in the object-to-know, making of itself the empirical object of its own knowledge, the distinction between non-empirical conditions of knowledge, and knowledge disappear into an illusionary unity called Man. This means that the subject *qua* not fully empirical element, within its doubling, disappears from the horizon of the questionable, annihilating the very possibility of every critique.

In order to know what the subject of knowledge is, it must necessarily appear in the empirical field (as object-to-know) and this appearing is, paradoxically, the disappearing of the transcendental subject, for by definition it is non-empirical.[12]

It is this disappearing of the transcendental conditions which traps critique. The trap, though, is hidden by the structure, a sort of orientation, which the analytic of finitude itself has given to modernity. The projects of modern emancipation, in fact, take for granted the research of a foothold that, wanting to reach a valid knowledge of its transcendental conditions in the subject, in history or in language reproduce the illusion just described.

According to Foucault, then, it is necessary to find a way to revitalize the transcendental dodging of the trap of Man. For this reason, he elaborates experimental ways to escape from the so called anthropological turn. And it is at this point that the Kant of the *Aufklärung* comes back.

[12] Foucault, *The Order of Things*, London, Tavistock Publication, 1970, Chapter 9.

According to Foucault, the Kantian answer concerning what *Aufklärung* is, marks the beginning of a new form of inquiry. In answering the question of what *Aufklärung* is*,* Kant is raising a question about what is present, thereby renouncing conventional approaches to the analysis of history. This Kant inquires about history without looking for any origin, internal finality, or for any eschatological realization. The reflection on history transforms itself into the problematization of the present with its characteristic form of normativity.[13]

Following Foucault's reading, this is the moment in which Kant relates himself as philosopher and his work to his own present, recognizing himself as necessarily belonging to a specific time. Kant situates his thought in the center of modernity, postulating the necessity of that time for his critical enterprise to work.

The present is not a fragment of the linear time, but a sort of enigmatic entity which brings together the fact of modernity and the duty to realize that modernity. In this history and in this narration of Kant's work, the human being is both object and actor.[14]

What Foucault sees is a movement: from the dimension of truth to the one of the being. If thought has a place and a time, then the critical task switches from the "analytics of truth" to the "ontology of ourselves", that is to say from the idea that reason comes to have of its own limits to the practice of questioning the limits brought about by the play between power and truth.[15] It is not a question about the concept of modernity but about modernity as self-inquiry, as a form of self-interrogation.

With and after Kant, modernity—that is to say, our present *qua* self-interrogation—has no longer the problem to understand its relation with the past or with a supposed

[13] About this point see *Dreyfus* and *Rabinow* (1986, pp. 109-122).
[14] Foucault 1984a, pp. 41-82.
[15] Ibid., p. 59.

novum, what matters is the *mode* of relation to the "today" and to the subject which enables this "mode of relationship".[16]

Kant is the threshold between the classical age and modernity. He, according to Foucault, stands for a sort of perpetual entrance into a critical relation to contemporaneity.[17] We, the moderns, still find ourselves standing in the shadow that this entrance casts, for, according to Foucault, the relevance of Kant's interrogation is not extinguished.

The issue of *Aufklärung* as both diagnosis and duty, as question concerning the involvement of thought into its own time, defines the philosophical perspective of both Kant and Foucault. It regards the philosophical exercise that for both characterizes itself as *sapere aude:* the courage to know the limits —however, not with the same aim in mind—, and the courage to identify the task of the present.

That mode of questioning the present, according to Foucault, drives Kant to recognize, or at least to gesture towards, a radical difference nestled within time. The present and with it *Aufklärung* are exactly this difference, or to use Kant's term the exit, the "way out" [Ausgang]. A "way out" from a state of immaturity that concerns, at the same time, the subject and its will at a personal level, but also the entire *Menschheit*. What Foucault suggests is that present is for Kant both a condition and a decision.[18]

The "immaturity", in fact, is something that does not depend on a sort of lack of rationality. It is rather the condition in which one finds oneself when one lacks courage to think outside of pre-constituted powers—scientific, religious and political power. And it is

[16] Ibid., p. 41.
[17] Foucault 1966a.
[18] Foucault 1984a, p. 35.

this ethical tonality that informs Kant's philosophical task and that Foucault recognizes as his own orientation.

In *Qu'est-ce que la critique? Critique et Aufklärung* Foucault rephrases Kant as stating that immaturity is the result of both an excess of authority, and of the already mentioned lack of courage. This means that the question that concerns the Foucauldian critique and the Kantian *Aufklärung* cannot be separated from the issue of "the art of not being governed or, better, the art of not being governed like that and at that cost".[19]

With these passages, what has been always seen as Kant's moderate proposal for a compromise with the imperial power, is rediscovered by Foucault as the very ground of a different critical thinking.

There is a manner of thinking — critique within the context of the *Aufklärung* — which is immediately also an ethical attitude. What is at stake with the question of critique is not a thought that shows and actualizes itself in an unlimited and spontaneous practice of insubordination. It is not the radical turmoil against every form of governmentalization, rather it is an exercise of freedom which knows (or at least tries to know) its own limits and tries to modify and overcome them in a sort of game of necessary co-presence with power.[20]

Critique is a form of antagonism that does not aim to replace or annihilate power. It aims to know and recognize power in all its articulations and transformations. The pastoral, disciplinary, and technical aspects that power has displayed over a millennium of the so-called western history needs to be scanned, as it were, within and without the subject. Critique's aim, according to Foucault, is to maintain high and constant the level of attention

[19] Foucault 1978a, pp. 44.

[20] Foucault often uses the concept of game. Even if there are not direct textual connections with some classical works about games, it seems plausible to think that he drew on the complexity and the richness of the concept as unpacked in the works of Bateson (2000) and Caillois (1961).

to ourselves with the awareness that we are placed within certain dispositives of knowledge and power which shape the way we feel and think but that, at the same time, as imperfect identities, make self-questioning a possibility for us.

> If governmentalization is [...] this movement through which individuals are subjugated in the reality of a social practice through mechanisms of power that adhere to a truth, well, then! I will say that critique is the movement by which the subject gives himself the right to question truth on its effects of power and question power on its discourses of truth.[21]

Critique places itself at the juncture between knowledge and power, and observes their inter-depending processes which have as effects the production of both spontaneous ways to obey the regimes of truth/power and the possibility to call into question those regimes and the self, which is partly involved in regimes themselves.

In this sense, the limit on which critique rests is critique itself. Critique acts as liberating and knowledgeable will when the problem of how to liberate oneself emerges as such. In order to experience and then oppose the given danger, it is not required to apply any normative theory or normative theorizing.[22]

Kant, with his way to evade the contents of the *Aufklärung,* maintains what, for Foucault, is fundamental: a sort of negative principle of thinking, an open address. Not just the study around the affirmative manifestation of knowledge and its forms (the analytics of

[21] Foucault 1978a, pp. 47.

[22] Even if Habermas sets the question of 'why fight at all?' in quasi-universalistic terms thus putting Foucault (the hypothetical one) into the impossibility to provide an answer from his standpoint, the question remains valid. This question can be understood as highlighting the fragility and the uncertainty of practices of liberations which are always taking place in a specific historical context. The practice of reassurance does not form part of the practices of liberation.

truth), but also a complex and not entirely clear interrogation of being—a question about a mode of existence. [23]

It is then through this reading of Kant that Foucault starts outlining the first elements of what, from then on, is an explicit ethics of existence. Critique, as autonomous practice of thinking– as art of existence – which carries out a distancing (opposed to the adherence of the governamentalization) from both the subjectification and the self, calls for a transformation of the webs of truth-power relations, of the world, and, above all, the self. It is a work on oneself which aims to transform the sovereignty of a self already known, thought and produced by the dynamics of truth-power, the always situated will to take care of that self without necessarily adopting any institutionalized knowledge.

But what does it mean to reason by oneself and to abandon the state of immaturity? The term Kant uses to answer this question and that Foucault reckons as essential is a term that Kant has employed in the *Kritik* too, namely *räsonnieren. Räsonnieren* designates not a generic use of reason but a specific one; it defines the activity of reasoning which has as its aim reasoning itself: "reason for reasoning's sake" in Foucault's words. It is this activity that allows reason to question its own limits, the conditions of its licit and proper functioning. This kind of intellectual activity is inherent in the unfolding of the transcendental critique, as presented by one of the two Kants which Foucault considers to be more a target than an interlocutor.

[23] Foucault introduces in an article from 1966, in which he comments on the French edition of Cassirer's *Die Philosophie der Aufklärung,* the concept of 'double nostalgia'. According to Foucault, Kant has trapped ('médusé la pensée occidentale') the modern thought within this duality. His double interrogation (about the forms of legitimate knowledge and a mode of existence) forces us either to look to the ancient Greeks, to whom we turn to ask about our relation to being (see, for instance, the works of Heidegger and Hölderlin), or we look closer to the *Aufklärung* in order to question the limits and the authority of knowledge (Marx and Lévi-Strauss). Nietzsche can be seen, according to Foucault, as the 'monstrosity' who plays on both the fields in this sort of small grand narration of the two last two centuries of so-called western philosophy. In Foucault 1966, pp. 546-547.

We are faced with one of those textual passages in which Foucault moves with extreme agility but that, once unpacked and carefully reconstructed, shows all the complexity of his thinking.[24] "I am not trying" says Foucault "to show the opposition that Kant would make between the analysis of *Aufklärung* and the critical project*".*[25] And actually he does not show it, also because according to him it would be easy to show the opposite, but he still constantly gestures toward that tension. Yet, he then goes on to state:

> But it no less remains that Kant affixed the understanding of knowledge to critique in its enterprise of desubjectification in relation to the game of power and truth, as a primordial task, as a prolegomena to any present and future *Aufklärung.*[26]

It is not the analytics of truth that opens the very possibility for a "the way out" towards an autonomous subject; on the contrary, it is *Aufklärung* as difference within the present which is essential to any critical movement of (self-)comprehension. Yet, there is a lingering of the Kant of the *Kritik* that peeps out with his twin as "the understanding of knowledge", even if understood as "desubjectification in relation to the game of power and truth", is still a "primordial task".

In order to focus on the possible way to distance oneself from the ties of power, Foucault is inevitably forced to face that very question at least in order to re-pose it in a post-transcendental context.

It is, in fact, an archeology of knowledge that Foucault will engage with before openly dealing with the question of being and thinking otherwise. Thinking and being

[24] It would be interesting to trace the rhetoric use Foucault makes of the argumentative inversion. I am confident that even from that point of view his writing would resemble that of ancient philosophy.
[25] Foucault 1984a, pp. 49.
[26] Ibid., p. 50.

otherwise, in fact, requires both Kants and both modernities, once in an antagonistic relation, and once as allies.

Coming back the term *räsonnieren,* it too follows this two-faced modernity. It is shaped by the authority of the *a priori* forms of truth, while, on the other hand, it interrogates that very authority by posing the question about the mode of being of the knowing and experiencing subject. And this is not another version of the analytics of finitude, as Han-Pile understands it, is because the present as condition of possibility remains open.[27]

This opening, on the one hand, explains how the quasi-transcendental and the empirical are not conflated; and, on the other, also suggests why Foucault is not in the position to give a plain and clear definition of the differences between truth as condition and truth as effect of practices, languages, and so on. And it is for this latter reason that Foucault, due to his intention to test the possibilities that appear only at the edges of the conditions of truth, cannot provide any theory (as he repeatedly says) but just a knowledgeable practice, so to speak. For theories work within the limits of their legitimacy not on the limits.[28]

However, saying that also this question of the reasoning for reasoning's sake belongs to the duality is not enough. It is must be explained how this necessary antagonism can be integrated into the economy of an ethics of the existence.

The thinking which is formed within its own limits and according to its licit functioning is an essential source for critique. However, Foucault also places himself in an antagonistic relation to the study of the possibility of legitimate knowledge, focusing

[27] See Han-Pile Foucault's Critical Project, Stanford, CA, Stanford University Press, 2002. . I will face Han-Pile's criticisms at the end of this chapter.

[28] At this regard is useful to think about the way Butler tries to develop this point in her work *Critique, Dissent, Disciplinarity,* that is starting from the distinction between a 'legitimate critique' and the 'critique of legitimacy'. (2012: 26)

instead on a mode of thinking which can form *on* the limits rather than within the limits. In addition, Foucault, still consistent with his critique of the analytic of finitude, completely, rejects the sovereignty of residual transcendental self as a possible foundation of his hyper-critique.[29]

Yet, this does not imply that Foucault can do without any concept of transcendence and its importance in analyzing the emergence of the enabling conditions of knowledge and (modes of) being.[30] Not only does this concept not disappear, but the structural contradictions of a historico-material transcendence become productive in Foucault's hands.

The analysis of both modernities will have to undergo the analysis of what makes knowledge possible and accepted in that way: archeology. The analytics of the historical transcendental conditions are necessary to the development of the work on the autonomous forms of the care of the self, even if the theoretical instruments for their comprehensions are different.

The study of the historical *a priori* is necessary but not sufficient. It is also essential to face the fact that the kind of thinking enabled by the order of discourse is, in turn, situated in a web of forces with which it interacts. The micro-physics of these forces is then another limit to investigate, and which then requires an appropriated tool: genealogy.

If philosophy can be an ethical practice, and as such be exercised as mode of critique, one cannot think the free and performative exercise of thought as a sudden, total, and immediate liberation. That would be a philosophical and theoretical paroxysm, as

[29] Deleuze feels so confident to write that 'Foucault's major achievement: the conversion of phenomenology into epistemology' in Deleuze 1988, p. 109.

[30] In this very regard Han-Pile holds that even the concept of 'regime of power-knowledge' is a transposition of the category of the transcendence, and that in Foucault's work there is the chance, at least, to raise the question about the possibility to read the whole analytics of power as a 'metaphysics of power'. See Han-Pile 2002, p. 9.

Foucault pointed out to Jean-Louis Bruch.[31] The bravery of the critical thought is to know what is knowable – as Foucault puts it when concluding his lecture of 1978.

With this goal in mind, Foucault's critique, similarly to Kantian *Aufklärung,* identifies and interweaves three thematic fields: truth (the knowledge which is accepted as truth), power, and ethics. That is to say, the classical triad: *aletheia*, *politeia* and *ethos.* His research does not aim to flatten or assimilate one to the other. On the contrary, the task that Foucault sets for himself is to analyze the specificities and the heterogeneities which allow for the existence of relations between these fields. It is a critique that, in virtue of its heteronomy—it is always critique *of*—can be an archeological, genealogical and ethical practice.

It is exactly this immanent and heteronomous character that makes critique an analytics which, by pointing at the non-fixed structure of reality, allows real possibilities to surface.

This ethical attitude which informs such a practice, then, transforms critique from a negative analysis of the limits, which one must recognize as not insurmountable, into the positive practice of possible overcoming. But again, this possible overcoming requires a systematic "historical investigation into the events that have led us to constitute ourselves and to recognize ourselves as subjects of what we are doing, thinking, saying."[32]

For this is necessary in order to find a strategy that may open up the possibility of a transformation.[33] Necessity which is prescriptive because it calls upon us to think possible solutions to unwanted and undesired conditions, but which does not presuppose a *theory* of what can legitimately be recognized as undesirable.

---

[31] Foucault 1978a, pp. 75-76.
[32] Foucault 1984a, p. 46.
[33] This passage is particularly complex. Foucault (1978a: 75-76) seems in some way to gesture towards the *will* of not being the subject of that thinking, saying, and doing as it were a foundation. Foucault denies this, saying that it is never an absolute will, but always a will of not being governed in a certain manner.

The absence of a theory does not mean that critique is a blind practice. In fact, the absence of theory is of capital importance to understand how the inseparable game of knowledge and power produces its effects within its webs of interactions.[34] These effects are, at the same time subjugating/subject-enabling and open to the indecision of the present – a present in which malaises and possibilities of liberations are both given as already *partially* determined.

Hyper-critique does not eliminate either power or knowledge, and has no ambition to do so; what it can do instead is avoid the tightening and ossification of their interactions into unquestioned knots, as can be seen in the case of the mental institution. In that case, for instance, the psychiatric knowledge and the actual rigid relations of power are (or risk to be) so strictly implemented to the point of erasing the very possibility of questioning their historical conditions of emergence. In other words, the form of psychiatric institution is regarded as a field of a (supposed) rational necessity.

In this section I have shown how Foucault's work on *Aufklärung* functions as both retrospective index, and general prospective on his entire oeuvre. That essay, in fact, renders it possible, provided the right lens, a reading of archeology, genealogy and care of the self as zipped together by an ethics of the existence, which itself is premised on the concept of critique. In the remainder of this chapter, I will show that Foucault outlines a feasible critical practice, but before defending the feasibility of his approach I will analyze the theoretical tools that are necessary to make decisions pertaining to not being governed in a certain manner and/or at certain costs. Such a reconstruction has the aim to show

[34] Foucault often stresses the difference between effect and product. The effect is neither something wanted (knowledge and power have no will) nor something causalistically determined and it is for this reason that effects are open to the transformation.

exactly the kind of specialized work these tools are required to implement in order to have a fully functioning hyper-critique.

The Archaeology of Knowledge and the Possibility to Think Otherwise

While a significant part of the modern critical project aims at grounding the validity of scientific statements, Foucault inaugurates his intellectual path by starting from the empirical fact that there is scientific knowledge.

Starting from this fact, he works backwards in order to look at the logical and historical passages that indicate that some statements and practices crossed the threshold of what is considered valid knowledge. Foucault calls investigation of such passages archaeology.

According to Foucault, knowledge works as historical *a priori,* that is to say, not as condition of validity, but as conditions of a practice of discursive thought.

Thought, in turn, is for Foucault the thing that places subject and object within reciprocal relations, and thought itself is influenced by that *a priori*. Consequently, knowledge effects both the knowing subjects and their field of action, by virtue of being for them both a historical enabling condition and limit.

The archaeological analysis of knowledge, then, has a double function. It is supposed to tell us something about the way we think, but by focusing on the historical passages that have actually allowed for the emergence of those forms of knowledge, it also reveals the contingent side of what is regarded as true.

The historical analyses of knowledge suggests that even if there are reasons—sometimes very good reasons—to think the way we do, those reasons are (often) not the result of a strictly necessary process, that is to say they do not equate to causes.

When the specific thoughts and the broader field of knowledge do not derive from strictly causally determined processes, then it means that there was in the past and that there is still in the present the real possibility of thinking differently. Archaeology, which investigates the knowledge which is off-stage points at those moments in history—the turning points—where it is still possible to see that knowledge could have taken other paths. In so doing archaeology shows that (some) things can still change as they do not rest on logical chains of events; instead they—the parts of knowledge which is not merely logically derived—rest on bare historical events. Archaeology then has the power to restore some fields of knowledge to their positivity (which is not synonymous with the arbitrary).

This instrument of investigation has, in Foucault's work, a prominent place, as even before he had an understanding of how to address a positive exercise of critique, he intended to explore the actual possibility of extracting oneself from undesirable forms of thinking.

Given this aim, the first intellectual move Foucault makes, probably not envisioning the subsequent ones, is to look for both the conditions that subjugate the subject and the possibility of counter-acting such a subjugation—taken together: a quasi-transcendental perspective.

It is the dispositive of the clinic and the field of medical knowledge which give the first opportunity for archaeology to take the floor.

Foucault, in *The Birth of Clinic*, starts both shaping and applying the archaeological method in relation to the medical glance and the rising institution of the hospital. In this

text, he traces the birth of a new scientific language, subtracting it from the accepted interpretation which recognizes in Bichat—the French anatomist—a sort of evolutionary turning point. According to this understanding, medicine comes to be as it is, in virtue of the corrections of previous mistakes and superstitions, only within this correcting process that an objective knowledge of both the body and of diseases could have been brought about.

In opposition to such a narration, since the very first pages, Foucault pushes the reader to face the empirical data at our disposal concerning what practitioners were actually doing starting from the second half of the XIII century. The immediate effect of reading the documents of that period is the discovery of a much more composite and less linear history of the medical science.[35]

What emerges from Foucault's study is that it was not, for instance, the withdrawal of religious believes that allowed for the development of anatomy. As he observes that "This reconstitution is historically false. Morgagni had no difficulty in the middle of the eighteenth century in carrying out his autopsies; nor did Hunter, some years later".[36] Foucault describes, instead, a change of the practices as something not immediately ascribable to a change in the manner death and diseases were understood. On the contrary, it is a series of (partly) contingent necessities imposed by the institution of the clinic that enabled some practices rather than others. It is, then, the institution of the clinic and the practices linked to it which function as "historical and concrete a priori of the modern medical gaze".[37]

[35] Another not irrelevant effect is the repugnance willingly caused by Foucault's descriptions as rhetorical expedient here and at the beginning of *Discipline and Punishment* to consign the reader to the alive matter he is describing

[36] Foucault, The Birth of Clinic, Routledge Classics, London, 2003, p. 153.

[37] Ibid., p. 237.

The study of the archive, as system of formation and transformation of the concrete medical statements, places Foucault in a position to attack a causal understanding of the emergence of scientific languages. For Foucault, in fact, the system of formation and transformation of scientific languages is not a kind of pre-existing field of rules that is able to capture the occurrence of an event or of a statement before or beyond it happens. On the contrary, he sees this functioning as connected to the strength of what has been said, and to the detailed report of what has happened.

In other words, there is order, and this order can be described and studied, but such order is not necessary; it is, on the contrary, the effect of historical events which once have been established, has, in turn, the ability to effect (not to cause).

This "analysis of the local discursivity" makes appear the rules—some rules—in their extreme fragility. A fragility that is not due to a time which infiltrates the systems from without, but that is partly constitutive of the system itself.

Archaeology, then, with its descriptive function, by making the positivity of the archive visible, makes also possible the mobilization of it as different form of knowledge, that is to say descriptive histories of regularities which do not amount to science.

As explained in *Archeology of Knowledge*, the system of the archive is a field, the general horizon where, on the basis of historical constrains, the statements appear in their irreducible positivity. Such a horizon and its constraints allow for the identification of the relations that the statements entails. In so doing it also makes possible the study of the dependencies which exist between the objects-to-know, concepts, modes of enunciations, institutions, economical and political practices, techniques and strategies.

This activity, by itself, according to Foucault, is able to expel the illusion of a general history from which it is possible to derive, retrospectively, a specific genetic knowledge.

Archaeology, explains Foucault, is not a way to question language and thought as such; on the contrary, it is a historico-critical exercise which has among its polemical targets exactly the kind of highly formalistic and illusory approaches which believe to be capable to know language and thought by drawing exclusively on them.

The strategy that archeology adopts against that illusion is to not presuppose any unitary generating principle, and, even less, any principle that is situated below or beyond the actual enunciations. This means that Foucault even while studying actual discourses, renounces (and actively apposes) any the idea of an interiority.

This very strategy, that for methodological reasons excludes the active function of the subject when the object of the study is the archive, could suggest a sort of immobilism. Yet this impression is misleading. In a provocative manner, Foucault, answering to those who hold that this kind of approaches could lead just to a conservative political approach, says:

> [D]iscourse is not life; its time is not yours; in it you won't reconcile yourself with death; […] in every sentence reigns the nameless law, the blank indifference.[38]

Although this seems to be just a sort of endorsement of the accusations to be a conservative thinker, Foucault holds that in order to foster a 'progressive' thought, it is not necessary to make of the subject "an universal operator of all the transformations". What he, with his archaeology can do, instead, is to identify "the different levels and functions which subjects can occupy in a domain which has its own rules of formation", and try to

[38] Foucault, "Politics and the Study of Discourse" in *The Foucault Effect*, University of Chicago, Chicago, IL 1991, p. 72.

understand "the manner in which the discourses in their positivity are part of a system of correlations with other practices".[39] This means that even if Foucault, when engaging with the tool of archaeology, stresses the importance and the role of anonymous systems without mentioning a clear ethical addressee, he, at the same time, by inserting these anonymous systems within a field of dispersion, multiplication, and differentiation, is indirectly subtracting the subject from the rigid constitution of her identity, showing, and, at the same time, opening up the possibility to be otherwise.

In a manner that is a trivialization of Foucault's work one could try to explain this implicit subtracting in the following way: Foucault shows through a meticulous analysis of the scientific enunciations that they are neither radically arbitrary nor univocally or causally determined. There are reasons for them to be the way they are, but the processes which form these reasons and the determinations of the interconnections between them are not as compelling as they may seem to be. Knowledge's webs, as seen from their archives, are neither so tight nor that onto-logically determined as they may appear. The ability to show this fragility is what allows Foucault to write in the introduction of *The Birth of Clinic* that "[t]he research that [he is] undertaking here therefore involves a project that is deliberately both historical and critical, in that it is concerned—outside all prescriptive intent—with determining the conditions of possibility of medical experience in modern times.[40]

Even if hyper-critique is not yet in place, the study of the archive shows itself, already at this stage, as one of its essential quasi-formal roots. This is because the individualization of the historical *a priori*, that is the singling out of regularities which enable the field of knowledge but not cause it, is the *conditio sine qua non* of the 'way out' of the *Aufklärung*.

[39] Ibid., p. 70
[40] Foucault, *The Birth of the Clinic*, p. xxi

In other words, a study that shows the enabling conditions knowledge, and, at the same time, exposes those conditions as partly belonging to the contingent, makes them appear (also) in their relative malleability.

The archaeology, as first moment of the ethical approach to the study of the objectification/subjectification, leaves itself open to a certain vulnerability. In fact, the objectifying power of the discursive orders coexists with, until almost blurring into, the opening of a space of antagonistic dispersion entrusted to the critical work of the archeology.

Yet, such an overlapping between objectification and dispersion is internal to the wider question of the singularization of being that, according to Foucault, is an ambivalent process of passive/active and productive/destructive elements. And it is this very ambivalence which shows the possibility of a critical work without providing any guarantees.

This means that, even if archaeology is an essential pawn in the strategic game of critique, the questioning of the subject—for questions of method and field of analysis—expunged from the analysis. Archaeology looks at the discursive and *anonymous* form of knowledge as what unfolds both the enabling condition of thought and also the possibility for transforming it.

If I have spoken of 'changing those relation' it is because I work with the complete work of Foucault in mind and because his use of the term critique pushes also the reader who is unaware of the following development to ask what is critical about the archaeological method.

In fact, what is critical about the archeological work is somehow what is not written on the pages of those texts but what they are able to bring about: the interstices of possibilities they open up but that they do not mention directly.

In order to give shape and voice to these openings Foucault needs a way, a conceptual junction, to link the question of knowledge with the one of subject. It was clear to Foucault that struggle and power were essential components in this process, but it took time before the concept of power acquired a less problematic form. It is only when the concept of governametality was clearly shaped that critique became an explicit topic: only at that point did Foucault have sufficient instruments at his disposal to give critique an actual chance.

The Power to Disconcert

The analysis of the historical constrains of knowledge especially after the publication of *The Order of Things* caused intense criticisms. Most of them were raised in the pages of Sartre's review *Les Temps Moderns*. Foucault was accused to elude history. Archaelogy was reckoned as insufficient to deal with the question of time, and completely unable to analyze the historical transformation from a diachronic point of view. The questions raised concern the ways the historical a priori come to be, and about their way and reasons of extinction.

In some way, such criticisms have contributed to the choice Foucault made to pass from the analysis of the history of knowledge to the one of power. It is exactly at moment in which the archaelogy is seen with suspicion because deemed unable to account for a sort

of fluidity of the historical process because freezing each historical moment into its fragmentation, that Foucault starts to think, that on the contrary, he has left too much space to history and to its paradoxical transcendental function. All what is not discursively calculable, in fact, has been so far neglected.

The archeology, if left alone, risks to ossifying the historical facts, that is to say the historical discourses, without being able to conceive of a more complete theoretical picture in which the singularity of the existence emerges from history without the justification of any historical a priori. In other words, and trying to employ a terminology which does not belong to Foucault, he is concerned about the possibility to get stuck between a scientific paradigm and another without being able, with the archeology alone, to grasp the grammar of the shift.

By solely describing the archive we are unable to account for the affirmative strength of a discourse, of knowledge. Archeology is unable to account for the ability of truth to constitute entire fields of objects and relations (prisons, hospitals and universities for instance). Putting it otherwise, truth has a power that the archive cannot describe because it is not gathered there, rather it unfolds in the movable and material webs of the present.

Furthermore the archive can say nothing about the historically unexplainable happening of the contingency—the "formidable materiality" as Foucault puts it. The event, its emergence [Entstehung] and its provenience [Herkunft] requires a sort of compensative analytics.

Genealogy will be the new instrument added to the 'tool-box', and the webs of power-truth the new notion which reveals a new front on which the difference within the present must be searched, and from now, even potentially cultivate.

The question of power, of the governmentalization, and of its normalizing/productive effects require an explicit examination of the subject qua subject which could not be part of the archeological enquiry. Power, dethroned by its juridical, repressive, and oppressive status, is made horizontal and it is entrusted to a reading that understands it as producing forms and as multiplied in moving webs of hostile and non-hostile forces. Power understood in this manner, according to Foucault, is recognizable, fully at work, and even established by, what he defines as "discourses of truth". This account of power is not only established and fed by the normalized and institutionalized knowledge, it requires, almost claims the production of truth – inquires, records it.

Such an imperfect—for unable to eliminate counter-functional and non-functional effects—circular relation between power and knowledge provides Foucault with a grid of intelligibility which opens up the possibility of new variety of historical incursions.

These incursions, and in particular the powerful and almost theatrical representation elaborated in *Discipline and Punish*, are meant to clearly display more than explain the passage from a hierarchical form of power to the modern normalizing one. The object of study of genealogy, as force directly and reciprocally implicated by knowledge and its truth, loses its repressive aura to become productive and intimately connected to the subject.

Despite the productive function of power, Foucault explains that in this case too his aim in not the shaping of theory of power, rather he intends to design an analytics. Power, then, is not the new name of an absolute order, rather it is both the cluster of the actual relational webs and a grid capable to maintain together an explanatory principle along with the mobility, in places absence of sense, the openness of the present, and even of what has been.

The genealogical practice does not need a normative framework to justify and ground its own evaluations and this not, as interpreters as Taylor and Fraser suggests, because Foucault wants to provide a neutral historical description. Instead, the absence of a theory of power derives from the fact that he is looking for a practice and such a practice cannot provide its own foundations if its aim is a liberating one. It can only search for some of its own limits and for the limits which disclose themselves within past and present;

> In reality, what I want to do, and here is difficulty of trying to do it, is to solve this problem: to work out an interpretation, a reading of a certain reality, which might be such that, on one hand, this interpretation could produce some effect of truth; and on the other hand, these effects of truth could become implements within possible struggles.[41]

What Foucault is saying here is that the genealogical study of power and its effects of truth, has (or can have) an immediate critical side. Genealogy is not the same as critique, but the detailed narration of the mechanisms in which the webs of truth-power set up is a new instrument to disclose a possible fertile space for critique.

The historical analysis, this time, is not guided by the research of what before has been defined as 'discontinuities', rather the genealogical form of enquiry looks for the "insurrection of subjugated knowledges" with the aim to reconstruct and disclose an "historical knowledge of struggles".

Now, beyond Foucault's rhetoric and emphasis which sometimes seem to contradict with the tone the strategic abandonment of a repressive notion of power, what here is relevant is the understanding of what he means by "subjugated knowledges". Power and

[41] Foucault, *'Clarification on the Question of Power', in Foucault Live: Collected Interviews, 1961-1984, ed. Sylvère Lotringer, Semiotext(e), New York,1996*, p. 189.

knowledge, it has been said, implicate each other; a certain configuration of the web of power is maintained by a certain kind of recognized knowledge which articulates itself in scientific discourses and in material institutions. This relation is neither pure nor automatic, and leaves open a myriad of competing narrations — what Foucault with an unhappy expression defines *le savoir des gens*.[42]

This kind of non-normalized knowledge is the space within which 'local' criticisms can and are produced. Foucault says that:

> What this essentially local character of criticism indicates in reality is an autonomous, non centralized kind of theoretical production, one that is to say whose validity is not dependent on the approval of the established régimes of thought.[43]

What is striking in this fragment of a lecture delivered at the beginning of 1976 is the incredible affinity with a passage of the first work Foucault elaborated two years later concerning the historical emergence of critique in the so-called western thought. The passage reads:

> When governing men was essentially a spiritual art […] not wanting to be governed was a certain way of refusing, challenging, limiting (say it as you like) ecclesial rule […] It meant questioning what sort of truth the Scripture told […] Let us say that critique is biblical, historically.[44]

[42] Foucault, 'Two Lectures' in *Critique and Power: Recasting the Foucault/Habermas Debate*, M. Kelly ed., Cambridge, MIT Press, 1994, p. 21.
[43] Ibid., p. 20.
[44] Foucault, 'What is Critique?', In *The Politics of Truth*, S. Lotringer ed., Semiotext(e), Los Angeles, 1997, p. 46.

It seems clear that in the middle of the interrogation of the mechanisms of power, the ethical orientations that concretizes its constant critical work on both sides – the one of the historic-theoretical comprehension and the one of the research of surmountable limits – is already functioning as guide of the genealogical researches.

The critical and locally emerged knowledge [savoir], antagonist to stronger and recognized forms of knowledge [connaissance], constitute the counter melody, so to speak, internal to an open and non-reducible material dialectic which is navigated through, and at some extended animated, exactly by those non-abstract criticisms.[45] What Foucault does with the work on critique is to openly recognize and fully elaborate this ethical kernel, finally opening the space to the first and more important effect of the interplay between normalized knowledge and the practices of normalization: the subject. A subject who is intimately constituted by power and truth, who has a history and an identity which emerges from the relations of power. Yet this subject is neither the same as his history nor equal to that identity that emerges from the interplay, and as such it is capable to transform the webs of power-knowledge and along with them also herself.

The discontinuities highlighted by archaelogy on a more diachronic level and the differences exposed and fostered by genealogy on the field of present (the 'today' of the Aufklärung), both sought as breach into the authorities of the regimes of truth, acquire a precise strategic place when seen from the point of view of critique.

Retrospective works always bear a degree of falsification, as it were, and I am not suggesting a twenty years detailed plan shaped since the beginning with precise intention to arrive, passing from the passive side of the subjectification, to the active one of

[45] "By *connaissance* I mean the relation of the subject to the object and the formal rules that govern it. *Savoir* refers to the conditions that are necessary in a particular period for this or that type of object to be given to *connaissance* and for this or that enunciation to be formulated" (Foucault 1969, p. 15 footnote 2).

transformation. On the contrary, it is the ethical intention – not a worked-out plan – which has guided this passage, which has made necessary, work after work, that kind of intellectual movement which once laid down into the concept of critique has appeared in its entire route.

And this constant ethico-critical movement is all the more evident when Foucault, through a sort of temporal transposition into the ancient Greek, finds the way to use his now complete 'tool-box'

The Care of the Self

Critique – it has been said – unfolds itself on a double level: on one hand, the necessary study of the historical conditions of the possibility of truth and of its actual functioning and, on the other hand, this same movement of comprehension reveals itself as *indocilité,* as a reasoned and reasoning practice which aims to modify the relations with one's self and with the world.

This idea to modify the relation to the self and the world without justifying it through legitimated forms of knowledge reaches its full elaboration during the series of lectures delivered at the Collège de France (1981-1982). In those lectures Foucault examines the techniques, half way between art and craft, of the *techne tou biou*. It is the art of existence according to which one should take care of one's self.

More than one concept, the care of the self is a cluster of concepts that Foucault clarifies in the course itself and in the last two volumes of *The History of Sexuality*. He, in a sort of constant practice of linguistic transgression, pulls these concepts out of a modern context of interpretation to restore them to an ancient use. The cluster of the arts of

existence contains concepts such as the ability to know and decide, competence and inventiveness, the mere patient work on the self and the one of ethical beauty. But let's proceed with order and hopefully a bit of clarity.

Since life [bios] has made its debut as synthetic concept which unfolds and rests under the modern conception of sovereignty, (personal and institutional) the practices of social and personal domestication, so to speak, have always presupposed (as existence and/or finality) the knowledge [connaissance] of that life and its forms as logically prior to each of the practices enacted to take care of it.[46]

Foucault, in order to open within this modern conception the possibility of an alternative relation to the self and to the other, seeks in history the moment when the link between institutionalized and formalized knowledge and the modes of relation to the existence were not yet settled.

The field of analysis at this point switches from the classical and modern periods to the ancient one. Texts as the *Alcibiades* and the *Laches* become the preferential objects of analysis. The two Platonic dialogues are exemplary of a period in which the care of the self [epimeleia heautou], was still swings between the knowledge of the self and soul, and the art of the care of the self, that is, the aesthetics of the existence.[47] Where esthetics is understood as a practico-critical attitude which escapes the knot truth-subject re-proposing it under a completely different meaning.

The question which Foucault tries to answer with those lecture was a question that he had already opened at the time of *The Birth of Clinic* and that along and after the itinerary of the analysis of the apparatus of power-knowledge, delineates itself in an

[46] About the shift see: Foucault, *The Birth of Bio-Politics,* New York, Palgrave Macmillan, 2008; and Foucault, *Society Must be Defended, New York, Picador, 2003.*
[47] Foucault 2001 [lectures of the 6th , 13th , and 20th of January 1982].

explicit way: how it is possible that the singularity produced and caught into the webs of subjectification can reappear, as it were, as an active subject?

It is Foucault himself that in *Subject and Power* summarizes his oeuvre saying that his whole work has always been about the "three modes of objectification which transform human beings into subjects".[48] The first mode objectifies the subject under the form of science looking at it as an entity which speaks (philology, linguistics, the studies around the general and rational grammar), works (economy), and lives (biology). The second mode regards the "dividing practices" (psychiatry for instance) which objectify by dividing the subject with and without the self, that is, the subject which results from those practices is estranged from her own self and from others. The last mode deals with both the way the subject comes to see and recognize herself as subject within a domain or forms of relation (sexuality, for instance) and the way she shapes herself as subject. And then going on, Foucault states: "thus it is not power but the subject, which is the general theme of my research". [49]

The genealogy of the modern subject must then be understood as a fragment of a wide critical history of knowledge and of its connections with the modes of being. This history, as Foucault explains in the introduction of *The Use of the Pleasure*, is an analysis of the "problematizations through which being offers itself to be, necessarily, thought — and the practices on the basis of which these problematizations are formed".[50] Such analysis, therefore, cannot avoid the study of the practices through which the aesthetics of existence gives itself.

[48] Foucault, 'The Subject and Power', in *Michel Foucault. Beyond Structuralism and Hermeneutics,* Chicago. University of Chicago Press, 1982b, 208.
[49] Ibid., pp.28-29.
[50] Foucault 1984b, p. 11

To analyze a subject in its moral form, it is then not sufficient to look at it as an agent, as the actor of its conducts, but it is necessary to see it as the effect of a double constant movement: subjectivation and askesis. *The Hermeneutic of the Subject* is therefore a necessary step in the analysis.

In this collection of lectures, Foucault analyzes the different articulations of the interweaving and constituting links between truth and subject. According to Foucault, the so-called western history of the relation subject-truth is dominated by a principle epitomized into the Delphic injunction *gnothi seauton* (know yourself). Such principle is the only way to access the subject and therefore truth.

Foucault holds that the interplay between subject and truth has not always constituted itself through knowledge [connoissance], but that it previously had its roots into the principle of care.

Care of the self and the knowledge [savoir] of that self have always been coexisting principles, but according to Foucault their relation has undergone both a hierarchical inversion and a progressive qualitative transformation due to the re-qualification of both the elements.

The motives that have led the principle of the care of the self, first to rise and, then to be disqualified in favor of the principle of knowledge until the reaching, for the latter, of what Foucault, with "a bad, purely conventional phrase", defines the 'Cartesian moment', are reconstructed starting from an observation of pure historical character.[51] The Delphic injunctions was not the philosophical foundation of the relation with both the self and the

[51] Foucault 2001, p. 14

others, rather it was a simple "counsel of prudence", a pragmatic invitation to consider one's means, character and ends before venturing into any enterprise.[52]

Subsequently, Foucault shows that the more complex ambitions of the knowledge of the self which will emerge through Plato's work find their "justificatory framework" only within the care of the self. The access to being is neither immediate nor self-evident, rather it pass through a complex cluster of practices, attitudes and processes. In addition, such art of the existence looks at the knowledge of the self not as aim but as mere stage of a rational behavior whose core is a knot of moral rationality:

> Having to care about oneself is not just a condition for gaining access to the philosophical life, in the strict and full sense of the term. You will see, I will try to show you, how generally speaking the principle that one must take care of oneself became the principle of all rational conduct in all forms of active life that would truly conform to the principle of moral rationality.[53]

The nexus between rationality and the practical behavior that is subjectivity and the possibility to access it truthfully, through the history of the *epimelia*, shows that that very subjectivity is (or at least can be) not immediately accessible. Rather it is the result of an attentive work. The subject and its truth are exactly this work made to enter the self without pre-supposing any self. The subject is neither the starting point of the relation nor its landing one; it is the relation itself. In the way I read Foucault it is clear that he is not describing here a logic of recognition in an idealist fashion as Han-Pile's analysis suggests. In her words: "the idea of subjectivation operates, by means of recognition followed by a

[52] In order to strengthen this starting point Foucault mentions W.H. 'Roscher, Weiteres über die Bedeutung des E zu Delphi und die übrigen grammata *delphika'*, in Philologus., Vol. 60 (1901), p. 81. (Foucault 2001, lecture of the 6th of January 1982).
[53] Foucault 2001, p. 9.

voluntary action of the self by the self, is afflicted as initio by and intellectualizing character that contradicts the letter as much as the spirit of the genealogical method and seems to revive at the heart of Foucault's analyses the idealism from which he had always wished to escape".[54]

Now, it is certainly true that Foucault, at that stage, deals with knowledge in a way that resonates as Hegelian. In the introduction to *The Use of Pleasure* , the verb 'to recognize' even if often placed side by side with the term 'experience' is certainly too central to be made unproblematic, as Gutting does, simply suggesting to not read it as the technical term of idealist philosophy.[55] Yet, I read it as the new Foucauldian attempt to unsettle the conventional use of the philosophical terminology. Foucault, at this stage, is trying to trace the moment in which knowing oneself was not equal to the activity to recognize oneself as object of knowledge, rather it was a different knowing process based (also) on the care and the productive experience of the relation to a non-objectified self. The care of the self, therefore, can be read as the precise attempt to reformulate (upsetting it) an internalized master/slave logic. Within that kind of logic both recognition and the transformative work of the self on the self are central as central they are in the later Foucault's works, but there is an element in the care of the self which completely differs. In the experience of recognition that Foucault describes, the transformative work does not pass through the necessary objectification of the transformed side of the subject (her body). In Foucault the spiritual transformation does not include an institutionalized and objectifying knowledge neither of the body nor of its conducts. The absence of the objectification, once more, emphasizes the transformative and practical side without surrounding to any intellectualizing character. On the contrary, the transformative practices

[54] Han-Pile 2002, p. 185.
[55] Gutting 2010, p. 33.

and the non-institutionalized knowledge linked to them, highlight the actual chance of the subject to place herself on the open limit of the present which is in turn the complex and mutable difference between the conditions (and the constrains) of truth and its effects.

What is at stake here is research of a critical practice that makes possible affirmative work on the limits of the subject and of the relations with the others in a way which may do both without knowledge [connoissance], and a legislating I. Foucault does not aim to an unfruitful negation of the concepts in themselves; the target is, rather, to endanger their unity and identity. Truth and subject are essential to any critical practice, but if critique is the repositioning of the limits of these two elements, then it must be possible to think their limits as not fixed.

One by one, each of the heralds of the western rationality which pretends to rest within their own limits are named in order to be transfigured. Truth, subject and even soul are narrated by Foucault following an alternative history which, in virtue of the *epimeleia,* provides him with the chance to open these concepts and their relations to the experimentally positive side of the critique.

The care as "general stand point", "way of behaving", as a "certain way of attending to what we think and what take place in our thought" and also as practices exercised by the self on the self through which one "takes responsibility for oneself and by which one changes […] transfigures oneself" , is the core of the ascetical work of critique.[56]

Foucault looks at Plato's dialogues to identify and rescue a practice of subjectivity as transformative relation to oneself and the world. In opposition to a true self which is sure into its self-evidence and rescued by the radical hypothesis that is being itself to be untrue, Foucault sets his attention on a process of subjectivization ascetico-spiritual that does not

[56] Foucault 2001, p.11.

presuppose any true self. Spirituality, derived from the Greek practices of the self, "postulates that truth is never given to the subject by right".[57] The problem of the subject that so clearly, according to Foucault is put in evidence into the *Alcibiades*, was not access to truth. The essential question which runs through the text is: 'How do I have to transfigure myself in order to produce a truthful relation to myself and the world?'

This means that the question of truth was immediately also an ethical question and that that dimension so essentially ethical of truth has ceased being so only when through the Christianization of the subject, truth and self reached, if not an identity, at least a direct and privileged link.

As I previously specified in this work, in a manner that can be really confusing, Foucault transfigures also all his terminology. Thus, when he looks in a favorable way to the pursuit of truth, one must bear clearly in mind what Foucault is speaking about before imagining a sort of conversion.

While modernity is characterized by an infinite tendency to indentify knowledge [connaissance] and truth, before Plato and with Plato (Aristotle is seen as the exception of the ancient age) truth is coextensive with the practices of care of the self. There is not truth beyond the practices themselves.

Obviously the question that Foucault has to face at this point is 'what is the self of which one has to take care? The soul [psukhe] – Foucault answers through Plato's words. It is not the concept of 'soul-substance' the one Plato employs, rather he (the Plato of the Alcibiades), discovers – says Foucault – the "soul-subject". That is to say, the non-instrumental relationship, as well as "position", at once "singular" and "transcendent" with

[57] Ibid., p. 15.

the other, with the world in general and with the oneself. And it is in virtue of the transcendence that the care cannot be equated either to medicine, economy or sexuality.[58]

Now if the object of the care is the 'soul-subject' the fact of the care must also require a form of knowledge. Such knowledge is not an introspective analysis rather an external gaze, as it were, referred to the divinity.

Foucault refers to this kind of mutual but contradictory necessity as "a dynamic entanglement, a reciprocal call for knowledge of the self and care of the self".[59] At least in the *Alcibiades*, ascetical transfiguration and rational principle of knowledge were still clearly interdependent even if caught in complex relation.

What has been the division between two modernities and two Kants elaborated in the work on *What is critique*?, is here, through the pages of the Plato's *Alcibiades*, traced back to a difficult coexistence which has more the aim to raise the same problem under a different historical prospective than to solve it.

Foucault's reading of Platonic paradox (the coexistence between the transformative and the comprehensive movement) shows, once again, the tension towards, and the use of an ethical perspective which necessarily, even if in different degrees, demands both the study of the limit and the practice of a possible overcoming.

Answer to the main criticisms

Foucault's work neither takes for granted that reason can reach an exhaustive knowledge of itself, nor that any degree of (self)knowledge is, by itself, liberating. On the contrary,

[58] Ibid., pp. 56-57.
[59] Ibid., p. 69.

Foucault's work provides a critical account of how and why forms of critique which try to found themselves on whichever normative/formal ground are doomed—for yet another version of the analytic of finitude. The analytic of finitude, as seen in the section regarding Foucault's archaeology, can be summarized as the impossibility for several kinds of radical critique to be sufficiently critical toward the piece(s) of knowledge they use as enabling condition(s).

Yet, Foucault does not confine his critical activity to the deconstruction of radical critique. He also tries to elaborate analytical tools able to both avoid the trap of using supposedly safe presuppositions, and embrace the modern ambition to reactivate a thought that is able to disentangle its fate, as much as it is possible, from the "the furor of power". This task, whenever it had an actual chance to be accomplished, cannot be achieved through a path that has consistently led to dead-ends; it must pass through a different manner of thinking. But is it really possible to think differently while still remaining critical?

Foucault's answer to this question is positive but not a given. It requires the interplay of all his analytical tools to successfully avoid "the captivating and Faustian convention" of radical critique, and formulate a pondered 'yes'.[60] Precisely because it is not a given, making sure there is the actual space for that kind of thinking is preliminary to any possible constructive theoretical move. For Foucault this meant restoring knowledge to his evenementiality—that is to move the point of observation from within what is valid to the local historical processes (the events) which led to the formation of those specific criteria of validity. This movement, for Foucault, has demanded a specific strategic decision: to keep

[60] It is the expression Tully uses to refer to the practice of bounding critique to a normative ground (what I define radical critique). Tully, 'Wittgenstein and Political Philosophy', in *The Grammar of Politics*, C. J. Heyes ed. Ithaca and London, Cornell University Press, 2003.

the critical investigations outside the solidified bastions of formal knowledge, and of the related normative grounds. It is, in fact, only from the historical investigation of how the modern criteria of scientific correctness have established themselves that the practice of hyper-critique can move on to make its first fundamental steps: i) verify the absence of a causal principle (not being confused with reasons) at the basis of the emergence of the human sciences, and in so doing, at least, shed light on the fact that the possibility of reasonably thinking outside those sciences cannot be excluded; ii) use the theoretical space that opens up to explore how to positively exercise a critique without normative foundations.

However, it is exactly this combined strategy of proceeding by looking for the conditions of possibly, while refuting any (formal or substantive) normative ground that has raised the main suspicion of hyper-critique as being self-defeating. To be precise, the accusation of having developed a self-defeating critical practice is the common conclusion of two different, but interconnected lines of criticisms.

On the one hand, Foucault would fail because hyper-critique is not even able to provide a convincing account of the distinction between the historical *a priori* and the empirical.[61] Foucault's critical practice, according to this line of argument, falls prey exactly to the logic which governs the analytic of finitude that he ascribes to radical critique. Therefore, his analytical tools, in particular the archaeology, reveal themselves as faulty and consequently unable to comply with hyper-critique's explicit goals. That is to say it cannot account for: a) the fact that there is observable order and that formal knowledge concerning the human sciences obeys to it, and b) the changes that these orders undergo

[61] The idea that there is space for testing Foucault against the phenomenological standard of the transcendental/empirical division is suggested, among the others, by Dreyfus, Han-Pile, and Kevin Thompson, Historicity and Transcendentality: Foucault, Cavaillès, and the Phenomenology of the Concept, History and Theory, 47 (Feb., 2008): 4. This last one does not conclude Foucault has failed against that standard.

over time. In short, hyper-critique would constantly face the impossibility to do what it was designed to do.

On the other hand, according to a line of argument that can be seen as a partial translation of the previous epistemological concern to a practical level, it is the lack of a divide (in this case intentional) between the normative and the positive that renders critique unable to justify its very practice. The alleged reduction of normative discourses to mere product of the complex interplays between formal knowledge and power that Foucault's genealogy appears to imply the annihilation of the normative — given his conception both of power and formal knowledge. And a critique without normative ground, it is claimed, not only loses its critical power, it also is unusable despite Foucault's historico-political insights.

In addition to these first two lines of criticisms which try to show the existence of substantial problems at the very base of the architecture of hyper-critique in order to prove its self-defeating character, there is a third line which challenges its efficacy on different grounds. This last heterogeneous family of criticisms stresses the inability of hyper-critique to be as critical as it should be.[62] For even if one suspends the scrutiny of Foucault's critical tools, or accepts them as non-self-defeating, Foucault's almost exclusive attention to an individual and aestheticized form of active resistance leaves it without the force to bring about the desired results. Hyper-critique, then, either ends up by involuntary increasing the atomism of our modern societies, or, in a more sympathetic reading, shows the lack of a satisfactory reflection on how to implement critique at the level of collective practices.

In this section of the work, I reject the criticism that Foucault is not able to distinguish between the historical *a priori* (or the conditions of possibility in general) and

[62] P. K. Dews 1989, McCarthy, 1991; and in a way that radically differs from the two previous works Allen 2008 and 2011.

the empirical by denying that we lose the efficacy of the *a priori* Foucault indentifies once the substantial distinction between the transcendental and the empirical is erased. What we lose, I maintain, is the illusion that the critique we are looking for in order to answer to the open demands of modernity can be a *science* of the limit—whether we understand by limit something enabling, or restraining, or the combination of both. Unsurprisingly, the other side of this loss is, as pointed out by the second line of criticisms I mentioned before, the concern about whether such a critical practice can be deemed to be critical at all.

Concerning this point, I demonstrate that hyper-critique not only has the sources to rebut these criticisms, I also show that the kind of criticisms that have been levelled look appealing only if one has in mind a specific understanding of critique: namely, radical critique, that is, the kind of critique that is uncritical towards its own presuppositions.

In the end, I will point out how a renewed relation to the self which does not pass through the objectifying knowledge of the self, is not necessarily a veer towards either atomism or a strong forms of individualism. It is a change, I claim, that can trigger local modifications also in the interpersonal and political sphere in a manner that cannot be easily foretold. With this I do not deny that those potential interpersonal relations need a further political and philosophical investigation, but I see this very necessity as a positive sign—the sign that Foucault successfully suggests new forms of inquires along with new effective critical tools which can and ought to be employed also in future investigations. Perhaps this result is more erratic than what Foucault's critics would hope for, but it is in line with the spirit of a critique that is constantly ready to call into question even its own practice.

On the alleged confusion of the empirical with the transcendental

One of the potentially more destructive criticisms against Foucault's critical project comes from Beatrice Han-Pile.

Han-Pile thesis develops around a single key. According to her, Foucault's critique is a project which rests on the attempt to formulate a modified transcendental edifice that is completely given in history.[63] In order to be critical, such a philosophical construction must be able to keep a well defined separation between the empirical, and the conditions of (historical) possibility of that empirical. According to her analysis, all the attempts that Foucault has made in order to implement this project ended up in replicating the irreducible aporia he intended to depart from: the analytic of finitude.

The confusion between the transcendental and the empirical originate, according to Han-Pile, from Foucault discovery that the empirical (the actual documented scientific statements) encompasses regularities that he claims to have also the status of *a priori*. Such empirical regularities that Foucault accurately describes cannot be said to regulate what they are intended to regulate. In other words, an empirical description of regularities, according to Han-Pile, cannot work as historical *a priori* for this would generate a regression (the dynamic of this regression will be explained later). Though this is exactly what Foucault does with his "happy positivism". Consequently, following Han-Pile's reading, Foucault's

[63] Han-Pile's work has been (together with Dreyfus's and Rabinow's) the forerunner of a now flourishing genre of secondary literature which reads in Foucault as a more or less successful phenomenologist. See for instance Johanna Oksala, *Foucault on Freedom* (Cambridge: Cambridge University Press, 2005), who argues that archaeology and genealogy share "with phenomenology the transcendental mode of questioning as opposed to a purely empirical study of the subject, but it does not share the methodological starting point in the subject"(104). Kevin Thompson, 'Historicity and Transcendentality: Foucault, Cavaillès, and the Phenomenology of the Concept', in History and Theory, 47 (Feb., 2008). Todd May, 'Foucault's Relation to Phenomenology', in Gary Gutting (ed.), *The Cambridge Companion to Foucault*. Second edition. (Cambridge: Cambridge University Press, 2005), Timothy Rayner, *Foucault's Heidegger: Philosophy and Transformative Experience* (NY: Continuum, 2007), and many papers collected in the Foucault-Heidegger volume by Alan Milchman and Alan Rosenberg, (eds.), *Foucault and Heidegger*: *Critical Encounters* (Minneapolis, MN: University of Minnesota Press, 2003).

critical project, against its own ambitions, is not able to account for the fact that "there is order" and that this order is what makes it possible for formal knowledge to achieve its status of formal knowledge. In addition, given the shifting nature and the contradictions at the core of the historical *a priori,* Foucault is unable to explain the changes that these orders undergo over time.

If Han-Pile is right then Foucault's ambition to build a work that is always both historical and critical (or critical insomuch as historical) cannot be underpinned. The critical side would reveal itself as a groundless claim, while the historical analysis, once left uncoupled from the ethico-philosophical intention, would shrink to a collection of unrelated historical documents.

Han-Pile, in my knowledge, is the one who has done the most in order to shed light on how relevant it is for Foucault's entire oeuvre to develop his own critical enterprise in a productive opposition to the theoretical trap of the analytic of finitude.[64] Yet in her analysis of Foucault's attempt to construct an alternative critical approach – which is able to deal with the "experience of order and its modes of being" in order "to rediscover on what basis knowledge and theory became possible"[65] – she appears to misconceive the reason why, as she put it, Foucault searched "his own solution for accounting for the temporal variations in the "experience of order""[66].

The way Han-Pile deals with the search for a feasible historical *a priori* is by placing this research within an alleged attempt to make of the archeology (in particular, but not only) a theory whose goal is to *account for* knowledge, in the sense of providing a historicized justification of knowledge. However, this an unfortunate and problematic

[64] For instance see also Han-Pile, 'The Analytic of Finitude and the History of Subjectivity' in Gary Gutting (ed.) *Cambridge Companion to Foucault*, Cambridge, 2007.
[65] Foucault, *The Archaeology of Knowledge,* London, Routledge, 2002, *from now on* AK
[66] Béatrice Han-Pile, *Foucault's Critical Project*, Stanford, Stanford University Press, 2002, p. 41.

perspective, for it entails that a substantial part of Foucault's practice of critique is obscured and marginalized. By looking at Foucault's archaeological work as a critique which exhausts itself in the search for the enabling conditions of knowledge, Han-Pile belittles the transgressive and transformative ambitions of his critical project. This focusing exclusively on the enabling conditions, in fact, leads to Han-Pile's reading of Foucault as a sort of transcendental idealist who sees in the practice of "knowing knowledge" the necessary and sufficient condition to be critical. This interpretation comes at a very high cost. In order to sustain her case, in fact, she needs i) to dismiss Foucault's declared intentions to not look for a conventional transcendental project since he deems this inconsistent with what he actually does, and ii) to present the works in which Foucault gives a mature account of what critique is as a substantial change of perspective.[67]

She is thus able to keep her interpretative perspective only by ascribing to Foucault's critique philosophical intentions and standards which either do not fully pertain to his critical project, or are the absolutized version of some of the present ones.

Now, there are works that stress these problematics in Han-Pile's interpretation of Foucault, and in particular with the general philosophical perspective she uses to read him.[68] Yet, given their preoccupation to save Foucault from a transcendental project, they have opened up the path to interpretations that are equally partial in that they tend to downplay the role of the historical *a priori* in Foucault's critique.[69]

Highlighting that hyper-critique is a form of historical critique rather than a transcendental one, has the merit to bring to the fore the fact that Foucault wanted to have a transformative impact on the present, but it underestimates how seriously the

[67] Ibid., pp. 73-77.
[68] Gutting 1979 for instance, makes of the first Foucault a historian with philosophical intentions but with no real philosophical method, or more recent interpretations as Kooppman's 2010.
[69] To be fair to both Gutting's and C. Koopman's work, they do not ignore the attention that Foucault devoted to the conditions of possibility, but they do underplay their role in Foucault's critique.

French philosopher has interrogated himself about the conditions and the possibilities of doing so. In addition, it is exactly the Foucauldian deconstruction of radical critique which shows that a mere historical analysis could easily be the umpteenth form of the analytic of finitude—Taylor is a clear example of this.

In other words, proving that Han-Pile's phenomenological reading of Foucault's project is inadequate is an important step to restore his critical practice to the path of an immanent critique, but it is not sufficient. In order to not discover at the bottom of our critical reasoning a sort of Hegelian trick—the trap in which "he stands, motionless, waiting for us"[70] as Foucault once put it—it is necessary also to pay attention to whether Foucault was actually able to shape a convincing historical *a priori* which is neither a formal *a priori* in disguise nor nothing. If that was not the case there would be the space for a teleological reading of history, and therefore the abortion of hyper-critique. For either *without* an open and plastic historical *a priori* or with a formal *a priori,* history can be seen as going exactly in the direction it must go—and this means that there is no space for being otherwise. On the one hand, the absence of relatively mobile conditions of possibility would entrust humanity to the necessity of what it is, and as a result, history will discover itself as nature. On the other, a formal *a priori,* which is by definition universal and necessary, will make of nature something originally rational. In the end the result is exactly the same: history is heading home.

If we do not want this conclusion, and I certainly do not, it is useful to follow Han-Pile in her close examination of Foucault's conception of the historical *a priori*. This is for me instrumental for proving that hyper-critique is endowed with a form of *a priori* that, under the functional point of view, does not get lost in positivity, and that therefore

[70] Foucault, *The Order of Discourse*, inaugural lecture at the Collège de France, Dec. 2, 1970 (1972, p. 235)

Foucault has actually carved out for his practice at least the chance to be critical while playing exclusively on the empirical ground.

The strange notion

In her work, Han-Pile shows in a very detailed manner that Foucault's entire project is underpinned by the concern of freeing critique from the vicious circle in which Kant has placed it. More precisely, Kant, according to Foucault, has first discovered the manner to place the human finitude at the ground of any possible knowledge without falling prey to the illusion that such a operation equates to the possibility of knowing that finitude, and then he was deflected from that fruitful path by asking the question *Was ist der Mensch?.*[71] This question put the newborn modernity in the anthropological slumber.

This anthropological sleep, expression which is used synonymously with the analytic of finitude is, according to Foucault, a prolific illusion that has given rise to the multitude of humanisms which, even if armed with the best intentions, cannot but reiterate the problems they intend to solve.

Foucault, as Han-Pile points out, moves from the awareness that in order to pull critique out of this impasse it is fundamental to have conditions of possibility freed from any variation of the theme of the analytic of finitude. Modernity, explains Foucault in *The Order of Things*, is characterized by conceiving the humanity "as a concrete epistemological positivity".[72] This conviction that humanity can be an object for knowledge to grasp generates a doubling: the doubling of the transcendental subject in its objectified and therefore knowable empirical version. This doubling, which Foucault respectively defines as

[71] Kant, *Logik* (Werke, ed. Cassirer, Band VIII) p. 343.
[72] Foucault, *Order of Thing*, London, Routledge, 2002, p. 315.

subject of knowledge, and object of knowledge, inaugurates the birth of Man. As soon as the subject doubles itself in the object-to-know, making of herself the empirical object of her knowledge, the distinction between non-empirical conditions of knowledge, and knowledge disappear into an illusionary unity called Man. This means that the subject *qua* not fully empirical element, within this doubling, disappears from the horizon of the questionable, annihilating the very possibility of every critique. In order to know what the subject of knowledge is, in fact, it must necessarily appear in the empirical field (as object-to-know) and this appearing is, paradoxically, the disappearing of the transcendental subject, for by definition it is non-empirical.

This paradoxical movement of appearing while disappearing and disappearing while appearing that the knowing subject performs in the attempt to know itself lies at the very bottom of what Foucault identifies as a the trap that keeps radical critique from being sufficiently critical, and so are all the human sciences which find the object of their analysis "already there". Habermas, for example, even while trying to avoid this trap by placing the structure of language (and not an idea of the subject) at the ground of his critical theory, has ended up using the criteria of validity that are part and parcel of language as both enabling condition and output of discourse ethics, thereby obliterating the chance to question the validity criterion itself.[73] That is to say another version of the mistake by which the transcendental not only thinks to be able to recognize itself in its objectified version, but also ends up making of that piece of (unverifiable) knowledge the legitimating standpoint of any possible critique.

[73] Tully, 'Wittgenstein and the Political Philosophy' in *The Grammar of Politics*, Ed. C. J. Heyes, Ithaca and London, Cornell University Press, 2003; and M. Poster, *Critical Theory and Poststructuralism*, Ithaca and London, Cornell University Press, 1989.

The modern knowledge concerning the human sciences pushes itself too far compared to where it can legitimately allow itself to go. Yet, this, for Foucault, is neither the sign that critique must be abandoned, nor the proof that in order to be critical it is sufficient, to step back within a more secure epistemological justification. Han-Pile suggests, instead, that Foucault intended to embrace exactly such a critical project with the only difference of opening it to history.

However, my reading of Foucault's critical project radically diverges from Han-Pile's on this point.

According to Han-Pile, the necessity to seek conditions of possibility that are neither a "decorative illusion"[74] nor the replication of the transcendental/empirical double means that, for Foucault, Kant's critical project needs to be fixed, as it were.

The way he, according to Han-Pile, tries to fix it is by transforming it into "idealist theses" in which "the conditions of possibility of knowledge are not homogeneous with the objects that they determine"[75]. Yet the endeavors Foucault produces to shape an *a priori* (or its variations: regimes of truth and episteme) that could activate such a corrected form of critique end up producing only partial definitions which, despite their incompleteness and variations, still prove "doomed to the "homunculous" sophism"; as they all seek "to explain one element by the others that in fact merely repeat its problematic characteristic".[76]

This fatal confusion between the conditions of possibility and the empirical which generates this endless regress on which hyper-critique supposedly rests is detectable following the several definitions of the *a priori*. And if somehow the confusions of the very

[74] Han-Pile, Reply to Gary Gutting's review of Foucault's Critical Project: Between the Transcendental and the Historical, p. 4.
[75] Han-Pile, 2002, p. 43.
[76] Ibid., p. 142.

early definitions are excusable because, by Foucault's own admission, at the time he wrote the first archaeological works he was engaged in the enterprise to "throw off the last anthropological constraints" and "in return, to reveal how these constraints could come about"[77], but he was doing it without a clear method. Then, when he, at last, came to write *The Archaeology of Knowledge,* his explicit intention was to provide a general and more refined articulation of such an enterprise, Han-Pile considers it appropriate to apply the strict criteria of analysis that a method requires.[78]

The results, according to her, are rather disappointing. The historical *a priori* is there defined as:

> not a condition of validity for judgements, but a condition of reality for statements. It is not a question of rediscovering what might legitimize an assertion, but of freeing the conditions of emergence of statements, the law of their coexistence with others, the specific form of their mode of being, the principles according to which they survive, become transformed, and disappear.[79]

This means, Han-Pile explains, that the historical *a priori* works as a principle of selection, that is "among the vast collection of possibilities offered by grammar and logic has the function of circumscribing a more restricted domain by defining the conditions of possibility of statements in their character as "things actually said".[80]

Such a historical *a priori,* in Foucault's conception, has the clear intent to account for the statements that have been uttered. Despite its being a "purely empirical figure", the historical *a priori* "must nonetheless involve", Han-Pile infers, "a specific kind of

[77] AK, p. 13.
[78] AK, p. 17.
[79] AK, p. 143.
[80] Han-Pile 2002, p. 64.

determination, which must be distinct from a causal determination—as otherwise the very possibility of archeology as the study of the statements at their own level would vanish".[81]

Han-Pile, following in this Dreyfus and Rabinow, notices that the nature of a principle able to be empirical or descriptive while being prescriptive or effective is, at least, ambiguous.

Archaeology is a method which is able to characterize statements on the ground of the regularities observable in the very linguistic practices.[82] Yet Foucault is not ready to stop his analysis at the level of the mere description, and this is because what is at stake is also the grasping of how the conditional possibilities themselves could come about.[83]

Foucault, this seems to be the criticism of both Han-Pile and Dreyfus, surreptitiously reintroduces a principle of causality by claiming for his method more than it is actually ready or even equipped to justify.

The fact that the discursive practices are said to be governed by rules is, within the methodological limits of the archaeology, highly problematic. Foucault, Dreyfus writes, "illegitimately hypostatized the observed formal regularities which describe discursive formations into conditions of these formations' existence".[84]

The convincing descriptions that Foucault has elaborated by analyzing madness, life, labor, and language itself put him in the position to licitly state that there are regular relations between discursive practices, and what for those practices counts as subject and/or object. However, saying that the mentioned scientific discourses obey or follow rules while also producing these regularities is something different. In that very moment we witness, according to Dreyfus, an indirect account of the causal power of the discursive

[81] Han-Pile, 2002, p. 65.
[82] AK p. 74.
[83] Dreyfus, and Rabinow, *Michel Foucault: Beyond Structuralism and Hermeneutics,* Chicago, University of Chicago Press, 1983, p. 82.
[84] Ibidem, p. 83.

formations. Therefore the historical *a priori* illegitimately claims for itself the principle of causal efficacy.[85]

This detected divergence between Foucault's programmatic refutation of the causal principle as part of the archaeology's explanatory instruments, and what he actually tries to do with his studies of discursive regularities is just one side of the problem according to Dreyfus and then Han-Pile. Simultaneously, and in a direct conjunction with that first piece of their analysis, they also deem Foucault's declared intention to appeal to a form of determination which is not causation as not feasible given his choice to leave the conscious subject on the backdrop of his archaeological investigations.

In order to be prescriptive and therefore efficacious, the rules for the discursive formations should be clear in the mind of people who formulate the scientific statements; but this is not what Foucault claims since he describes his task in *The Archeology of Knowledge* as bringing to light the enabling conditions of the statements which have been actually formulated without appealing to the consciousness of the speaking subject. If this power cannot be located in the mind of the subject, then it must be the case that it is present in the discursive practices themselves, that is to say, the very place where the regularities have been discovered.[86]

This makes the enabling criteria a "strange notion of regularities which regulate themselves"—an invalidating circularity from which, according to Dreyfus, archaeology cannot free itself even while considering the complex manner with which language operates, as I will report in a moment.[87]

[85] Ibid.
[86] Ibid., p. 84
[87] Ibid.

This criticism not only finds the approval of Han-Pile who restates it, but the logical structure of such a criticism reveals to be useful in the analysis of the many levels of Foucault's work. Han-Pile finds, in fact, that explaining by means of what needs explanation can be frequently encountered in Foucault's works. It, in fact, can be found in all of the analytical tools he developed over the years.

The alleged failure of finding a form of determination which is not equal to causation leads to the irreconcilable dichotomy between descriptive value and efficacy of the historical *a priori.*[88] This shortcoming of Foucault's archaeology leaves hyper-critique unable to free itself from the logic of the anthropological double, showing instead the enduring of "a tension between the historical and the *a priori".*[89]

The inability to decide between the merely empirical, and an accomplished transcendentalism, according to Han-Pile, keeps Foucault's project stuck. He can neither support his intention to shape a viable historical condition of possibility for the emersion of formal knowledge, nor can he account for the changes in those conditions over time.[90]

According to my reading of the alleged descriptive/prescriptive dichotomy, and appealing to my interpretation of the whole of Foucault's critical project, a historical *a priori* such as the one conceived in the archaeology while intentionally turning its back to causal power, does not lose efficacy. If critique is understood as an informed practice rather than a theory, that is to say, a critique which does not claim for the discovered historical *a priori* a foundational role, then this *a priori* can be seen as sharing the same nature as that which it regulates, but not the same function.

[88] Han-Pile, 2002, p. 65.
[89] Han-Pile, 2002, p. 196.
[90] Ibid.

As Han-Pile rightly points out, Foucault needs a kind of determination that is different from the causal one. In addition, this kind of determination must be of the kind that can work even while not being consciously present in the mind of the subject.[91]

As Gutting noticed in his review of Han-Pile's relevant work, complying with rules that we are not conscious of is something that, as speakers, we all do. Obviously it does not mean that one cannot engage with such rules in a reflective mode, but this engagement is not strictly necessary in order to successfully engage, for instance, in complex linguistic practices.[92]

The simple fact that we, at least in some circumstances, obey linguistic norms in such a manner, does not tell us if this kind of obedience is prompted by a form of causation or something else. It is not obvious, in fact, if the rules which govern language have a necessary form inscribed in the brain as Chomsky suggests, or if they are parts of the linguistic practice itself.

According to Dreyfus, even if one finds the latter Wittgensteinian reading more convincing, it is still the case we are not allowed to say that these rules that we use while engaging in linguistic practices actually govern behavior—for instance the adoption of formalism as a way towards the scientific validation.[93] In other words, even in a Wittgensteinian interpretative framework, rules can be seen as self-sustaining, but not as self-regulating; the regulating part must rest somewhere else, namely in the social practices.[94]

---

[91] Han-Pile, 2002, pp. 65-66.

[92] Gutting, (2003, paragraph 8)

[93] Dreyfus, and Rabinow, *Michel Foucault: Beyond Structuralism and Hermeneutics,* Chicago, University of Chicago Press, 1983, p. 82.

[94] Here the full quotation: "But Foucault is not satisfied to accept social practices as a level of explanation. *Social regularities,* it *seems, require* a *further account.* In *linguistics two* alternative models have been proposed to *explain* how grammatical rules govern the formation of sentences. Either, like Chomsky and Lévi-Strauss, one can hold that grammatical rules are formal rules which govern practice by being instantiated in

The conclusions that Dreyfus draws from this analysis, and that Han-Pile subsequently employs to prove the inefficacy of the historical *a priori,* appear, in my view, to be derived from an unjustified mix of two separate issues, and therefore illicitly deduced. He overlaps the question of the efficacy of the rules at a given time with the emergence of the rules themselves.

Now, concerning the emergence of the rules, the criticism was one of the first that Foucault's archaeological works faced, and, at that time, he simply did not know how to solve it, hence the necessity to add genealogy to his analytical tool-box. However, this issue that Foucault recognizes and addresses in his following works, is simply not relevant in the analysis of the suspected confusion between the transcendental and the empirical (efficacy vs description). For what we are looking for is 'simply' a determination which is not causation. The full explanation of how that determination comes about is something that we can bracket, at least for now.

However, the mere fact that social practices are the ultimate explanation for the emergence of the historical *a priori* itself*,* can be stated without any fear of contradiction even within the degree of understanding of the conditions of possibility that Foucault presents in *The Archaeology of Knowledge.*[95] For if the *how* was at that stage a problem, the mere fact of the relative mobility of the historical *a priori* was not.

Since the time of the archaeological studies, in fact, Foucault held that "the a priori of positivities is not only the system of a temporal dispersion; it is itself a transformable

---

the brain, or else, like Heidegger and Wittgenstein, one can hold that the linguistic practices themselves have the power to sustain and perpetuate norms […]. In neither case one can say that the rules themselves actually "govern", "operate", "determine," or "limit" behavior. It is either the neurons or the social practices that have the causal power".

[95] It is worth pointing out once more that the lack of the genealogical instruments of analysis is a problem in order to give a full account of the dynamic of the emergence of a regime of truth, but not in order to explain the efficacy of the *a priori* that is in place.

group".[96] And this transformable group is not understood as an "empty figure that irrupted one day on the surface of time, that exercised over men's thought a tyranny that none could escape".[97] The transformation of the rules happens along with what is regulated "and if they are not modified with the least of them, they modify them, and are transformed with them into certain decisive thresholds".[98]

The fact that the rules are a conventional effect (which is by no means synonymous with arbitrary or voluntary) of the social practices is not a problem for Foucault; on the contrary, the identification of the *a priori* as a system of relations which is such "because it makes constant use of this group of relations" is part of Foucault's strategy for disqualifying the principle of causality in the historical understanding of the emergence of the human sciences and their objects.[99]

The conventional character of the rules is not a way to make of the *a priori* something radically accidental, and therefore equal to no *a priori* at all, rather something being a matter of convention, here amounts to being exposed to neither the conscious will of the individual, nor to the whims of the case. All this means that rules can be abandoned, and that they are indeed the result of social practices. Or, as Wittgenstein explained while illustrating what he meant by "the inherited background against which I distinguish between true and false", this kind of antecedent mastery can be compared to the bed of a river which is comparatively much more stable than the water flux that runs on it, but is still transformable (OC §§ 94-96).

This very peculiarity of language—that it can be both conventional but not shaped by the simple will of the speaker(s)—suggests that the conventional character of the

[96] AK, p. 144.
[97] Ibid.
[98] Ibid.
[99] AK, p. 59.

historical *a priori* rules out neither the self-regulating ability, nor the fact that these kinds of rules actually govern behavior. And this is true, pace Dreyfus, particularly in a Wittgensteinian reading of linguistic rules.

This is not an isolated overlap between Wittegnestenian and Foucauldian thoughts. In the last years a substantial and heterogeneous group of intellectuals has started to explore the assonances between Foucault and Wittgenstein, identifying some important analogies that have been then used to illuminate and support a Foucauldian critical practice.[100]

Ironically enough, Wittgenstein too has been accused of both suggesting yet another transcendental project[101], and, at the same time, of being uncritical exactly because he has rendered unavailable any transcendental standpoint.[102]

Tully, for instance, in an attempt to show the contradictions of a purely hermeneutical understanding of critique, addresses a common mistake that is here repeated by Dreyfus; that is to say the conflation of interpretation and understanding.

In order to illuminate the difference between interpretation and understanding, Tully relies on the Wittgenstein of the *Philosophical Investigations*. There Wittgenstein explains how, as a sort of mental habit, we are led to consider the following of a rule as involving the interpretation of such rule, that is to say, the necessity to enter in a reflective relation with said rule (PI § 201). Yet if this were true, we would be closed in an inescapable

[100] English speaking authors like J. Tully, C. J. Heyes, D. Owen, and L.M.G. Zerilli have experimented with Wittgenstein's concepts as instruments of political and critical analysis. In the pages of the volume that collects all those works (The Grammar of Politics, C. J. Heyes ed. Cornell University Press, London, 2003) Wittgenstein's and Foucault' analytical instruments are intertwined in order to both deconstruct a certain understanding of critique, and open up new possible scenarios. On the other shore, in a work edited by Gros (Frédéric Gros & Arnold I. Davidson (eds.), Foucault, Wittgenstein: de possibles rencontres, Ed. Kimé, Paris, 2011) the papers T. Tisala and L. Paltrinieri build comparative analysis between Wittgenstein's grammar and Foucault's historical *a priori*.

[101] See P.M.S. Hacker's defense of Wittgenstein (2013, p. 50).

[102] J. C. Nyìri, 'Wittgenstein's Later Work in Relation to Conservatism', in *Wittgenstein and its Time,* ed. B. McGuinnes, Oxford, Blackwell, 1981, p. 44.

circle of interpretations, for the interpretation of a sign is another sign (PI § 201). In addition, if understanding and interpretation were interchangeable, an interpretation would be able to determine the correct use of a sign as understanding does. However, it is not a reflective decoding of the sign that determines the correct use of it (PI § 85). But then what kind of understanding is an understanding that is able to build "the connection between a sign or a rule and its use"?[103] Wittgenstein answers by qualifying this kind of understanding as a form of "training", the unmediated ability to "grasp":

> It can be seen that there is a misunderstanding here from the mere fact that in the course of our argument we give one interpretation after another: as if each one contended us at least for a moment, until we thought of yet another standing behind it. What this shows is that there is a way of grasping a rule which is not an interpretation, but which is exhibited in what we call "obeying the rule" and "going against it" in actual case. (PI §201)

Therefore, Tully goes on paraphrasing Wittgenstein, if understanding is conceived in this manner, there is no real gap between the rules and their use: most of the time the practice is a non-mediated understanding, a non-reflective obedience. This does not mean that what functions as rules cannot be interpreted or described, but this happens when there is some problem with the understanding (which, as Tully notices, makes the reflective mode conditional on understanding).

This form of grasping itself, then, is not yet another interpretation standing behind an interpretation, and as praxis it can be both minded and unreflective.

Now Dreyfus is obviously not making exactly the mistake of not recognizing that we do use language in such a way, what he is saying is that the mere fact that we reproduce

[103] Tully, 2003, p. 37.

and sustain rules (also) by using them unreflectively is equal neither to bring about those rules, nor to account for them—which is both true and irrelevant here. What is instead relevant to the research of regularities that can have a regulative, but not causal, power is that if we put together the fact that languages are conventional, and that their use can be seen as unreflective obedience to rules (conditional but not causal) then the "strange notion of regularities which regulate themselves" does not look so strange after all.[104]

Foucault has shown with his archaeology that discourse formation has a degree of independence from the subject (which is a way to qualify the relation, and not a manner to misrecognize or ossify it), and that this way of working constitutes a conditionality to be taken into account while trying to be critical.

The autonomous functioning of discursive practices, once rescued from a causal reading, even while not being at the immediate disposal of the knowing subject's will—for the subject itself is conditional on, but not determined by discursive practices—can still be the object of historical comparative studies.

The fact that there is order, at this point, can be re-discovered. Given the theoretical space earned by introducing a conditionality which is not a causal determination, the observable order can be seen as a sign of the conditioned field of dispersion whose rules can be locally discovered through the comparative study of the regularities themselves.[105]

That "field of the dispersion" that is a conditioned form of the possible to which one cannot ascribe any unity (archive or episteme) is something that cannot be grasped and described in its entirety. "The archive cannot be described in its totality" writes Foucault

[104] H. L. Dreyfus, P. Rabinow, *Michel Foucault: Beyond Structuralism and Hermeneutics,* Chicago, University of Chicago Press, 1983, p. 84.

[105] Foucault never explained in a systematic way what he meant by 'dispersion' but inferring from the use he made of the term, a field of dispersion is a field of possibility which is conditioned but not ascribable to any unity.

with a sentence unusually direct.[106] For "it is from within these rules that we speak". But we can work locally. We can discover situated conditions which in a given time make a statement possible. This kind of study opens up the chance to be critical exactly of those conditions of possibility, as it can be the synthetic concept of life, for instance.

In a manner that has some analogies with how Wittgenstein distinguished between grammatical rules and empirical propositions, Foucault, with his local studies, can still discriminate between what works as conditions of possibility and the positive, but he cannot use those conditions as foundation, for this would be an illicit move. The degree of self-knowledge that it would involve is something from which a knowing subject is precluded from, a subject that is partly enabled by the rules she wants to know. It would be equal to the illusion of knowing our conditionality: the mistake that in its various forms is at the ground of the analytic of finitude.

This is not a mistake Foucault reproduces. His is not a theory, and regularities and *a priori* are not deduced one another. He is "not inferring the analysis of discursive formations from a definition of statements that would serve as a basis; nor [is he] inferring the nature of statements from what discursive formations are, as one was able to abstract them from this or that description".[107] The distinctions and the relations are established in a correlative manner. As a consequence of this, the distinction between a historical *a priori* and positivity is possible but it is clearly not a categorical division that is able to maintain or

[106] AK, p. 146.

[107] AK, p. 128. In this connection, Han-Pile quotes a fragment of the AK that could make my reading problematic. I report here both the quotation she uses and the one I used. But I avoid to actually engage with her analysis because Foucault will then go on with no theory. "The analysis of the statement and that of the formation are established correlatively. When the time finally comes to found a theory, it will have to define a deductive order." AK, p. 130; "But one can also see that I am not developing here a theory, in the strict sense of the term: the deduction, on the basis of a number of axioms, of an abstract model applicable to an indefinite number of empirical descriptions. If such an edifice were ever possible, the time for it has certainly not yet arrived. I am not inferring the analysis of discursive formations from a definition of statements that would serve as a basis; nor am I inferring the nature of statements from what discursive formations are, as one was able to abstract them from this or that description"; AK, p. 128.

establish the non homogeneity of the two. This is not only impossible for Foucault' critical practice, it is also not desirable. That would be the very sign that Foucault has abandon a critical practice in favor of a critical theory which inexorably loses its grip on what calls for critical scrutiny every time the scrutiny takes as a given either its object or its subject.

This means that Han-Pile is right about Foucault not being able to keep the historical *a priori* and the empirical sharply divided but this is a structural problem only within the kind of transcendental project she has ascribed to him. Yet, her reading reveals its partiality as soon as one sees that Foucault neither tacitly reintroduces the category of causality, nor fails to give a convincing account of the efficacy of conditionality— a notion to which I will come back in the next section.

Obviously this methodological decision of simultaneously renouncing to the *substantial* division between conditions of possibility and the empirical, and the knowing of those elements in scientific terms, that is to say, providing an epistemological justification, has a price: we lose the possibility of having a critique that can be a science. What we earn in return is the ability to investigate the captivating force of knowledge (or at least some pieces of knowledge) itself—a practice of the limit.

As Owen noticed, highlighting another point of contact between Wittgenstein and Foucault, it is important not to be forgetful of the fact that what keeps us captive rests beyond the true/false dichotomy. The picture that holds us captive, in the very famous passage under whose aegis Owen places his reinterpretation of Foucault's critique, is neither true nor false, but this does not diminish its power, on the contrary.[108] What works as condition of possibility—the picture—influences even our ability to question, and by

[108] Owen, 'Genealogy as Perspicuous Representation', in *The Grammar of Politics*, C. J. Heyes ed. Ithaca and London, Cornell University Press, 2003. The quotation is: "a picture held us captive. And we could not get outside it, for it lay in our language and language seemed to repeat it to us inexorably" *PI*

considering it as fixed and totalizing transcendental, we take away the possibility of reflecting on the very source of an insidious form of captivation.

The possibility to put in practice critical and therefore liberating strategies passes also through a philosophical practice willing to embrace a change of perspective from which it is possible to see that, as Tully puts it, "the ongoing activity of questioning and arguing in accordance with" certain standards is exactly what keeps those standards out the field of questionable.[109]

Empirical Insight and Normative Confusion

Now that it is has become apparent that Foucault with his archeological studies has actually opened up the chance to be critical by discovering a kind of conditionality which is not causal and can by studied with a comparative historical method, I can focus on the second line of criticism which regards the eclipse of foundations deriving from the swap causality/conditionality as entailing the impossibility of critique.

As soon as what is true and verifiable about the human sciences is shown in its reasonable fragility, that is to say, as strongly linked to their historical emergence, also the normative ground of the practical sphere starts to shake.

The strange notion of regularities which regulate themselves not only is not so strange, but it also allows Foucault to make a further step. Once he has rules that are neither necessarily conscious nor causal, and yet efficacious, he can focus on both the captivating force of those webs of rules, and the possible strategies to escape that force.

[109] J. Tully, 'Wittgenstein and Political Philosophy' in *The Grammar of Politics*, C. J. Heyes ed. Ithaca and London, Cornell University Press, 2003, p. 33.

This further passage, though, required what Koopman defined archaeology-plus, that is to say, the addition of a new analytical instrument: genealogy.[110]

While archaeology alone was able to account for the fact that there is order and that this order works and reproduces itself in a manner to place the knowing subject within conditions that are both enabling and limiting, the very micro-physics of these processes was not clear yet. The moment in which Foucault started to question the non-discursive force that made those processes an alive dynamic, power, and with it the subject, entered the scene, turning hyper-critique into something both more effective, and more difficult to digest especially from the perspective of proponents of radical critique.

The opening of the possibility for a new form of critique, from the perspective of radical critique, has coincided with a paradoxical loss. While the historical analyses highlight something relevant about modern sciences, institutions, and the capillary form of modern power, the very philosophical imports derived from those analyses are seen as not activating critique, but rather as hindering it.

The omnipresence of power and its interplay with regimes of truth that affect the subject and her possibility to be autonomous are seen as incompatible with the liberating claims of hyper-critique.

The idea that, borrowing Foucault's words, "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" eliminates even the mere logical possibility of reaching a fragment of knowledge, or a sphere of life foreign to power.[111] By implication, every form of normative foundation is precluded to hyper-critique. For, among other reasons, the ability of power to

[110] Koopman, 'Historical Critique or Transcendental Critique' in *Foucault Studies*, 2010, p. 110.
[111] Foucault, *Discipline and Punishment*, p. 27.

infiltrate even normative discourses makes them unsuitable as foundation. Hyper-critique, it is claimed, is therefore unable to justify itself, and autonomy as a form of critical practice has no place in such a frame.

The basic form of this criticism is voiced by Habermas who sees in Foucault's genealogical works a sort of metaphysics of power which has as automatic consequence the identification of validity and power and therefore the annihilation of every form of critique.

The relation that Foucault tries to articulate between bodies of knowledge and power is read by Habermas as a causal knot. Power in his reading is a sort of First Cause while all the rest, knowledge and subjectivity included, are, at the same time, determinations and extensions of that cause. Not only, according to Habermas, "are truth claims confined to the discourses within which they arise; they exhaust their entire significance in the functional contribution they make to the self – maintenance of a given totality of discourse".[112] That is to say, the meaning of validity consists in the power effects they have.

Obviously, if Foucault ended up making of power and knowledge merely the conditioned field of emergence and an effect, while, instead, trying to set up a description of how knowledge and power always interact, then Habermas would be right.[113] Within Foucault's analytics of power there would not be space for any critique, no matter what

[112] Habermas, *The Philosophical Discourse of Modernity.* Cambridge, MIT Press, 1987, p. 279.

[113] If what Foucault ended up actually doing can be the object of an investigation, his intentions are quite clear and there is little room for interpretation: [If I had] made [power and knowledge] identical, I don't see why I would have taken the trouble to show the different relations between them. […] What I set out to show was how certain forms of power that were of the same type could give rise to bodies of knowledge that were extremely different both in their object and in their structure. [...] We have, then, power structures, fairly closely related institutional forms – psychiatric confinement, medical hospitalization – that are bound up with different forms of knowledge, between which it is possible to draw up a system of relations based not on cause and effect, still less on identity, but on conditions." (*Politics, Philosophy, Culture*, 264 – 265)

kind of critique. In addition, the mere attempt to be critical would be a sort of parody staged on the behalf of the only possible subject: power itself.

Such a reading of Foucault's analytics of power has been widely disproven, and even sympathetic authors as Fraser and Allen have found Habermas guilty of a grave misinterpretation.[114] In different manners, both Fraser and Allen have shown how Foucault's understanding of power as capillary and omnipresent does not equate to the claim that power is everything, or that it is the only modality of human interaction. On the contrary, it helps to illuminate how certain forms of subjugation are particularly difficult to address with the instruments of radical critique.

Foucault's works show that discursive and non-discursive practices constitute the conditions of possibility for subjectivity to emerge, and that those conditions can be empirically and locally analyzed. The fact of the conditionality of the subject in itself is in no way an obstacle to either the ability to critically engage power and knowledge, or to critical self-transformation.[115]

Despite this sort of rehabilitation of Foucault, the absence of a normative ground is still deemed as highly problematic.

According to Fraser, the main difficulty lies in the fact that due to the absence of a normative platform, Foucault is not able to give an account of "why ought domination be resisted?".[116] She recognizes in Foucault's works a prolific understanding of modern power, and she goes as far as to acknowledge that the mere account of a domesticating power alone potentially is able to rule out some inadequate political orientations.

---

[114] Among the non-sympathetic see for instance Kelly, M. 'Foucault, subjectivity and technologies of the self', in *C*. *Falzon*, *T*. *O'Leary* and *J*. *Sawiki* (*eds*), A *Companion* to *Foucault*, *Chichester*: *Blackwell*. 2013.
[115] On this point see Allen 2011.
[116] Fraser 1989, p. 29.

For instance, Fraser points out that Foucault's understanding of power as a disciplinary force calls for a "politics of everyday life" that from a feminist point of view is much more useful and effective than Habermas'. And this is exactly because Habermas has a too narrow understanding of power.[117] Yet, this method of analysis which has useful outputs, according to her, under the philosophical point of view, has also limits.

Genealogy, in fact, works with "the suspension of the standard modern liberal framework which distinguishes between the legitimate and illegitimate exercise of power".[118] Nonetheless, Foucault's very intellectual practice clearly shows to be "politically engaged". This engagement, according to Fraser, tacitly appeals to concepts such as 'rights' or 'limits' which belong exactly to the suspended framework.[119] This contradiction, for Fraser, necessitates a solution that Foucault does not provide. The question about what is bad about power or at least about certain forms of power must be addressed, otherwise all what genealogy would produce is "a curious amalgam of amoral militaristic descriptions".[120]

Developing further her analysis of Foucault's conception of disciplinary power, Fraser identifies the sources of this, for her, puzzling theoretical position in a genetic fallacy that according to her affects the argumentative structure of the genealogical works.[121]

If Foucault's rejection of the liberal framework is to be understood not simply as a strategic choice, then it must be the case that he considers the very framework to be intrinsically undesirable. Obviously this is a normative judgment—a judgment Foucault cannot justify. Yet the analysis of disciplinary societies and the notion of autonomy that derives from them are considered, according to Fraser's reading, in negative terms by

[117] Habermas has addressed these criticisms, recognizing at least partially the validity of Fraser's analysis.
[118] Fraser, 1989, p. 18.
[119] Ibid., p.19.
[120] Ibid., p.31.
[121] Fraser, 1989, p. 50.

Foucault, for power partly places their possibility to be. The source of autonomy, however, cannot be captured in these terms. As Fraser points out, the mere fact that autonomy has no other way to come about but through some form of authority (as even Habermas recognizes) is, in itself, not a licit reason to judge it as unreliable. By performing this sort of faulty tacit judgment Foucault invalidated the normative standpoint from which he speaks, leaving himself without any resources to ground his reaction against a panopticized society.[122]

This analysis, though, rests on erroneous assumptions. Foucault's analysis of power, as Allen rightly points out, has an explicit starting point that Fraser is distorting.[123] "My point is not that everything is bad", clearly writes Foucault, "but that everything is dangerous, which is not exactly the same as bad".[124]

Now, even if this invalidates Fraser's criticism, according to Allen, it does so by moving the possibility for the genetic fallacy on a different level of Foucault's argumentation. The justification of the critical practice, thus, retains its centrality.

In the same way as Fraser holds on to Habermas' concern about normative foundations while correcting the argumentation, Allen, in turn, corrects Fraser's criticism only to make of her underlying concern something even more specific, and therefore more to the point.

According to Allen, Foucault has the merit to show how entangled autonomy and power are. This simple fact makes clear that critical theory must interrogate itself further about the processes of socializations, the role power plays within them, and the possible

[122] Ibid., p. 53.
[123] Allen, E., 'The Entanglement of Power and Validity: Foucault and Critical Theory' in *Foucault and Philosophy*, O' Leary, Farzon (eds.), Chichester, Wiley-Blackwell, 2010 , p. 86
[124] Foucault (1997b: 256)

manner to take all this in account in a productive manner.[125] Now, despite the fact that Foucault is not slipping into the genetic fallacy Fraser attributed to him, he still does not provide any help to positively develop his own insights about power. On the contrary, his almost exclusive attention to the relation the subject has with herself, risks turning the elaboration of the needed further philosophical investigations into something problematic.[126]

Allen points out that the fact that autonomy is dangerous (but not inherently bad) because historically formed, and because power forms part of its possibility, is still a normative assessment, and it is then this claim that is endangered by the genetic fallacy. However this potential problem, Allen suggests, can be avoided by drawing on Foucauldian philosophical sources.

A Foucauldian critique would result in a genetic fallacy only if it were unable to say anything about the "*causal history*" that brought about the form of autonomy under scrutiny.[127] For instance, if a person developed her own autonomy in a dystopian society, then considering that form of autonomy as unreliable would not invite any fallacy because, in that case, it would be the process which is under scrutiny and not the origin in itself. If we find a way to judge that causal history, that is to say, to discriminate between a licit historical path and one that is too much interwoven with power, then we can redeem hyper-critique from the possibility to commit a genetic fallacy, according to Allen.

She envisions two possible ways to formulate the needed judgment: internalist and externalist.[128] The first approach leaves the subject free to decide by herself. The subject in that scenario would reflect on her history and judge whether her autonomy is reliable given

[125] Allen 2008.
[126] Ibid.
[127] The italics as in the original Allen's text.
[128] Allen, 'The Entanglement of Power and Validity: Foucault and Critical Theory', in *Foucault and Philosophy* Edited by T. O'Leary & C. Falzon, Blackwell, 2010, p. 90.

the way she has grown into it, or not. The problem that arises immediately within this possible scenario is that, given the position of the subject, which is internal to her own autonomy, she is herself hardly reliable.[129] For the "agent's willingness to endorse the historical conditions may itself be the product of pernicious, autonomy-inhibiting influences".[130]

The second option, according to Allen, seems more promising even if not lacking of potential problems. What is needed according to this second strategy are "agent-neutral criteria" in order to judge the way autonomy has been brought about.[131] Establishing these criteria in a manner that is "neither arbitrary, capricious, nor ideological" is certainly a difficult challenge, but it is here that Allen believes we can use Foucault's own sources to solve this problem.[132] A "frank acknowledgment of the necessary entanglement of power and validity provides at least a partial solution to this problem" according to her.[133]

Now, in spite of the fact that Allen has voiced at different occasions that by using the knowledge concerning the entanglement of power and validity we can indeed build the needed criteria, she has not yet gone so far as to enter into the details of how such a construction would work, but I will come back to this point later.[134]

In my reading, Allen's analysis of Foucault's notion of power is the most accurate of the three mentioned in this chapter. It is extremely useful in order to make clear that Foucault's analytics of power does not rule out the possibility of critique, and that, on the contrary, it, as critique of critique, makes a precious contribution to improve upon critical

[129] Ibid.
[130] Ibid.
[131] Ibid., p. 91.
[132] Ibid.
[133] Ibid.
[134] For instance see Allen 2008.

theory.[135] Yet, it is also puzzling. It is puzzling not only for the way she intends to carry on Foucault's critical practice, but also for what concerns the underplaying of the very specificity of the form of critique he suggests.

Her concerns about how to employ Foucault's insights are licit, but the path she takes concerning the validity criteria seems to move all the problems Foucault himself recognizes in radical critique from one level to another.

In her monograph from 2008, where she had the chance to explore extensively what are the questions left open by Foucault's critique and what can count as a possible solution to them, she frames an idea of validity which is certainly interesting, but not devoid of problems. She suggests a sort of hybridization.

Foucault, in her account, lacks a theory of the public sphere; Habermas, on the contrary, has a fully fledged theory of it. With the right adjustments, Allen claims, Habermas' could amend Foucault's lack. Her idea of adjustment is to use Foucault's understanding of both power and history to situate validity. In this manner, political actors would come to an elaboration of validity criteria taking in account the fact that those criteria can never fully transcend the context in which they have been elaborated. The resulting criteria can then be used to support political action.

Leaving on the side the fact that the contextualization of the said criteria seems to merely turn Habermas into Taylor without addressing any of the problematic consequences this entails; the problem is not that such a notion of validity is per se incompatible with the Foucault's own form critique; it is rather that validity should not be considered as an upgrade of hyper-critique, rather it is a completely different critical tool.

[135] Allen, 'Foucault and the Politics of our Selves', in *History of the Human Sciences* 24: 4 (October 2011): 43-60

A form of valid critique cannot be radical, and, as we have seen with the two previous chapters, when it tries to give itself such a goal, it ends up invalidating itself. Hyper-critique, instead, is a practice conceived to be used only when a specific notion or practice related to the way we think and are needs to be reconsidered. For this reason a merger between a critique based on criteria of validity and hyper-critique would hinder the positive and specific critical elements both approaches can offer.

In my reading of Foucault, in fact, there is nothing that prevents a political engagement based on some forms of validity. As long as the system of norms in which one lives does not give rise to distinct negative experiences and/or questions that cannot be adequately addressed with the means made available by a form of critique drawing on already established validity criteria, one can precede entrusting practical judgment to this legitimated form of critique. Nothing makes this practice uncritical in itself.

The fact that those criteria have a context, and historical roots fed by multiple forms of power, as both Fraser and Allen point out, is in itself not an obstacle to autonomy. On the contrary, it just means that autonomy is always conditioned. These conditions have been revealed by Foucault's studies as both limiting and enabling conditions of our way of being and thinking. But he never tried to reject or qualify as bad enabling conditions *tout court*.

However, one has to recognize that when the regime of truth on which we rely, and the criteria that form it themselves produce effects that are experienced as negative, and which cannot be addressed by recurring to already valid criteria, then the form of critique which is necessary in those cases cannot rest on those very problematic criteria.

Without being able to simply take recourse to extent criteria, for they are the ones which is necessary to call into question, nor new ones, as there are cases in which there are

not yet alternative criteria at our disposal, the needed form of critique can be grounded, according to Foucault, only in a local will.

One could say that when it comes to make sense of hyper-critique in normative terms, we can conceive it as a form or normative negativism that legitimizes, in principles, all critical activities that are directed at overcoming aspects of the world that are experienced as problematic or negative. What is entailed is a form of autonomy of hyper-critique, in the sense that the negativity of the experience requires nothing else to render the engagement in hyper-critical activities legitimated.

In a passage of the *Groundwork* that Allen herself quotes, Kant writes that autonomy is "the property of the will by which it is a law to itself" (Kant, 1997: 47). Now Foucault is a surprisingly orthodox Kantian (sometimes). The will to oppose what is experienced as negative and cannot be adequately addressed drawing on the extent practices of justification, because they are internally linked with generating this problem, is the foundation of hyper-critique.

The subject does not decide on the ground of some validity criteria that a political, a domestic or a personal revolution is necessary. In some way that will is already decided by itself, in the sense that it is a negatively charged response to a particular context. In this specific sense, the local will Foucault speaks about finds its ultimate cause in itself.

Embracing that will and engaging in a transformation that enacts liberating effects, and consequently also a possible production of meaning and forms of knowledge [sovoir] is something different from embracing a set of valid criteria [connaissance].[136]

[136] The difference between sovoir and connaissance is a difference I have, in some form, dealt with in Chapter II. And it is worth recalling here that Taylor, according to my reading, is not capable to explain how the passage from savoir (meaningful) to connaissance (authoritative) could be seen as a homogeneous continuum (epistemic gain).

This means that while there is no inconsistency concerning the unmixed coexistence between a practice witch adheres to validity criteria, and one which is conceived to intervene when those criteria constitute a problem, in order to think a possible interactions between the two, one has to recognize their specificity.

And the specificity of hyper-critique rests exactly in its refractoriness when it comes to being turned into a theory (either epistemological or political) that is to say a valid set of notions and criteria. Such specificity is what allows hyper-critique to hold on to its goal of being both emancipating in character and capable to potentially call into question the legitimacy of what we know and are without falling prey to the analytic of finitude—the dead-end Habermas and Taylor face because they intended to build a radically critical *theory*, that is to say a form of critique that in order to be both legitimate and radically critical, entrusts itself to a piece of knowledge which is indispensable but that, contrary to their ambitions, can never be itself scrutinized.

It is then only focusing on this specificity that we can possibly start imagining political notions and practices capable to employ the richness of both hyper-critique, and forms of critique which are legitimated but that cannot call into questions their own aims.

Obviously, the recognition of the differences is also recognition of the different and specific difficulties and limits. My distinction would, in fact, hardly satisfy Allen's concerns about how to employ Foucault's insights without renouncing to some firm theoretical stand.

The fact that Foucault is successful in regards to its own aims, does not mean that even his very success does not raise further questions to investigate, or concerns such as how we know that the will we want to give shape to is not a dystrophic effect, or how we

determine that the transformations enacted are good transformations, so to speak, for instance.

My view on those concerns that Foucault's critical practice raise is that, on the one hand, he has some instruments to address, at least, some of them, and that, on the other, hyper-critique cannot be demanded to perform according to ends that are not its own.

As Foucault notices "[p]eople do revolt; that is a fact".[137] The will to revolt enters the scene as effect of a concrete constellation which gives rise to experiences of negativity that cannot be adequately addressed within the extent way of living and thinking. Starting from that local will, Foucault provides us with analytical tools and strategies shaped to critically engage, and even channel that will.

Once the experiences of negativity are there, they cannot be ignored. They are a symptom, and as symptom those experiences—while not justifying by themselves whatever kind of transformation which might solve or alleviate them—do call for both the understanding of their sources, and possible transformations.[138]

When those experiences and the correlated will to avoid them emerge, archeology, genealogy and care of the self provide interpretative grids, detailed historical accounts, and a repertoire of notions potentially capable to restore the present to its ontological fragility, or as Foucault would say, to its evenementiality. This means that when as subjects we start to experience a certain notion or practice as alienating, for instance, we can first of all start to understand the degree of necessitation of that notion or practice: why has it emerged in first place? What are the reasons which kept it in use until today? Why do I rely on it? If the

[137] Foucault, "Useless to revolt?", in Essential Works of Foucault 1954-1984,Volume 3: Power Ed. J D Faubion (New Press, New York), 2000, p. 451.

[138] The necessity to address the symptoms stands even when the negative experiences are identified as the product of notions and institutions one would qualify as progressive. The presence of the symptoms alone calls for investigations that may or may not result in an actual transformation, but that, in any case, are both compelling (see the normative negativism mentioned in this section), and, in principle, capable to allow for a further understanding of the object investigated.

answers to these questions reveal volatile elements—reasons that no longer stand, or that came about in the first place because of mere contingent justifications which nonetheless today are considered authoritative—we can try to loosen the grip of those elements and transform what is transformable.

However, these sources that hyper-critique has in order to address potential concerns—sources Foucault has performatively demonstrated as useful, while working on relatively narrow notions such as the clinic, sexuality, the prison—do not provide answers to all the concerns I have mentioned above. The reason why it does not provide those answers is that, in large part, hyper-critique cannot provide them. For instance, we cannot eliminate the danger that the outputs of the practice of hyper-critique will themselves bring about other and different negative experiences.

Yet, in spite of its limits, hyper-critique, as piece-meal practice, does not leave us completely unequipped. As in the case of both its normative and motivational grounding, hyper-critique has sources that can help us to navigate the outputs of its own functioning.

While calling into question the authoritativeness of a given notion in the light of specific experiences of negativity and finding alternatives to it is its aim, hyper-critique has not the ambition or the instruments to call into question our whole form of life (this would be yet another form of the analytic of finitude). Therefore, in order to evaluate its results we can appeal to a composite strategy. First, we can investigate whether a particular expression of hyper-critique undermines other valid criteria which we had no reason to call into question (and a positive result would call for further questioning). Subsequently, we can focus on the transformative effects of hyper-critique, trying to understand whether and in what way the experiences of the negativity have been alleviated, or explained in a

manner that while convincing does not comply (yet) with what is recognized as valid, or addressed in any other way a radically critical practice allows for.

The refractoriness in being turned into a theory that I have mentioned in the previous paragraphs may lead to the conclusion that, coming to the third line of criticism, the positive side of hyper-critique is engulfed in an aestheticized notion of resistance.

Only a subject able to establish a new relation with her self, truth, and power can be able uncover the mechanisms which led modern reason to the furor of power. The fact that resistance cannot but move to the care (as opposed to the knowledge) of the self is not necessarily a "reinforcement of social tendencies towards atomization" as Dews suggests.[139] True, when it comes to such practices there is always room for a failure as "revolution always risk falling back into the old rut", but, in the meanwhile, the opening inaugurated by the practice of hyper-critique brings with itself a much richer supply of virtuality, of real possibilities.[140]

As subjectivity is something intrinsically social for Foucault, the transformations of relations to self can trigger political transformations starting, for instance, from "the politics of everyday life" which Fraser mentions as prolific implication of the genealogical studies.

In addition, Foucault, in a work that once again is the re-elaboration of Kant's historical reflections, writes that "what is important in the Revolution is not the Revolution itself, it is what happens in the heads of those who do not participate in it or, in any case, are not its principal actors".[141] The study on the revolution that for Kant is the French revolution is used by Foucault to reinforce his understanding of critique as *Aufklärung* and to add that that permanent possibility for an exit can generate what Kant defines

[139] Dews, 1989, p. 40.
[140] Foucault, "What is Revolution" in *The Politics of Truth,* ed. Lotringer, New York, Semiotext(e), 1997, p. 93.
[141] Ibid., p. 91.

"sympathy of aspiration". The ethical dimension of the care of the self can lead to normative outputs and therefore to new moral codes (I use here the distinction ethics/morality Habermas employs), but the new moral codes themselves will be no longer a form of hyper-critique.

The *pars construens* of hyper-critique is given by the experimenting with new ways of thinking and being; nothing prevents it from being collective too, but consensus about the criteria of validity presupposes already both the field of discussion and the objects about which one is going to deliberate.

There are important practical problems with Foucault exactly because *Aufklärung* as exit cannot constitute a critical theory, yet the acceptance of the risk (*sapere audere*) is a necessary condition for being critical in that manner.

For the first time (among the authors I have selected) the hypothesis is put forward that critique, at least the kind of critique we are looking for, is not something that can be fastened to a piece of knowledge considered solid enough in order to turn it into a secure mechanism.

On the contrary, following Foucault one must embrace a certain fragility as general paradigm. Critique in his work appears as a never ending process that is ready to face the fact that what today has liberated me tomorrow can be the very cause of oppression.

Perhaps this constant task, this constant being on alert and building informed tools for practicing hyper-critique is the maximum we can ask for when what is needed is calling into question the assumptions of a discourse or of an institution. This may seem a small achievement compared to the endeavor to restore a foothold for reason, but it might be also the case that the task of radical critique is a sort of (very modern) battle against windmills that distract us from the philosophical battles that are worth engaging with.

The fact that "Foucault wants to give up one set of dangers for another" can actually be a wise choice when the first set of dangers keeps us from facing a set of dangers which, by contrast to the first one, also entails a promise to be liberating, and as such is worth facing.[142]

[142] Dreyfus and Rabinow, *Michel Foucault: Beyond Structuralism and Hermeneutics*, Chicago, IL, University of Chicago Press, 1983, p. 264.

Bibliography

Primary Texts:

M. Foucault, (1961) The History of Madness, London, Routledge, 2006.

---- *(1963) The Birth of the Clinic*, Pantheon Books, New York, 1973.

---- (1966a) *Order of Things*, London, Tavistock Publications, 1970.

---- (1966b) *Maurice Blanchot: The Thought from Outside*, Urzone, New York, 1987.

---- (1966c) 'Une histoire restée muette', in *Dits et écrits*, Paris, Gallimard, 1994, pp. 546-547.

---- *(1968)* 'Politics and the Study of Discourse', in *The Foucault Effect: Studies in Governmentality*, C. Gordon tr., G. Burchell, C. Gordon, P. Miller eds., Chicago, Chicago university press, 1991, p. 54-55.

---- (1969) *The Archaeology of Knowledge*, London, Tavistock Publications, 1972.

---- (1970) 'The Order of Discourse', trans. Ian McLeod in R. Young (ed.), *Untying the Text: A Poststructuralist Reader*, R. Young ed., Boston and London, Routledge, Kegan and Paul, 1981.

---- (1971) 'Nietzsche, Genealogy, History'. In Language, Counter-Memory, Practice: Selected Essays and Interviews, edited by D. F. Bouchard, Cornell University Press, New York, 1977.

---- (1974) 'Prisons et asiles dans le mécanisme du pouvoir' in *Dits et Ecrits*, II. Paris, Gallimard, 1994.

---- (1975) *Discipline & Punish: The Birth of the Prison*, Allen Lane, London, 1977.

---- (1976a) 'Two Lecture' in *Critique and Power: Recasting the Foucault/Habermas Debate*, M. Kelly ed., Cambridge, MIT Press, 1994.

---- (1976b) *The History of Sexuality* R. Hurley tr., New York, Pantheon Books, Vol. I, 1978.

---- (1978a) 'What is Critique?', In *The Politics of Truth*, S. Lotringer ed., Semiotext(e), Los Angeles, 1997.

---- (1978b) 'Clarification on the Question of Power', in *Foucault Live: Collected Interviews, 1961-1984*, ed. Sylvère Lotringer, Semiotext(e), New York,1996.

---- (1997) *Society Must Be Defended: Lectures at the College de France*, 1975-1976, New York, Picador, 2003.

---- (1982) 'The Subject and Power', in *Michel Foucault. Beyond Structuralism and Hermeneutics,* H. Dreyfus and P. Rabinow eds., Chicago. University of Chicago Press, 1983[2].

---- (1984a) 'What is Enlightenment?' in *The Foucault Reader,* P. Rabinow ed*.,* New York, Pantheon Books, 1984.

---- (1984b) *The History of Sexuality*. R. Hurley tr., New York, Pantheon Books, Vol. II (The Use of Pleasure), 1985.

---- (1984c) *The History of Sexuality* R. Hurley tr., New York, Pantheon Books, Vol. III, 1986.

---- (2000)``Useless to revolt?'', in *Essential Works of Foucault* 1954 ^1984,Volume 3: Power Ed. J D Faubion (New Press, New York) pp 449-453.

---- (2001) *The Hermeneutics of the Subject: Lectures at the College de France 1981-1982*, ed. F. Gros, tr. G. Burchell, New York, Palgrave Macmillan, 2005.

--- (2004a) *Security, Territory, Population: Lectures at the College de France 1977-1978*, Basingstoke, Palgrave Macmillan, 2007.

---- (2004b) *The Birth of Biopolitics: Lectures at the College de France, 1978-1979,* New York, Palgrave Macmillan, 2008.

Habermas, J. 1976. *Legitimation Crisis*, trans. T. McCarthy, London : Heinemann. [1973]

---- *1984, 1987.Theory of Communicative Action*, vols I and II, trans. T. McCarthy. Boston, MA: Beacon Press. [1981]

---- *1987. The Philosophical Discourse of Modernity. Twelve Lectures*, trans. F. Lawrence, Cambridge, MA: MIT Press. [1985]

---- 1990. *Moral Consciousness and Communicative Action*, trans. C. Lenhardt and S. Weber Nicholsen. Cambridge, MA: MIT Press. [1983]

---- *1992. Postmetaphysical Thinking. Philosophical Essays*, trans. W. M. Hohengarten. Cambridge, MA: MIT Press. [1988]

---- 1996*. Between Facts and Norms. Contributions to a Discourse Theory of Law and Democracy*, trans. W. Rehg, Cambridge: Polity Press [1992]

----- 2002*. Religion and Rationality. Essays on Reason, God, and Modernity,* ed. and intro. E. Mendieta, Cambridge: Polity Press.

----- 2003. "Faith and Knowledge" in *The Future of Human Nature*. trans.W. Rehg, M. Pensky, H. Beister, Cambridge: Polity Press. [2001]

---- 2003. *Truth and Justification*. ed. and trans. B. Fultner, Cambridge: Polity Press. [1999]

----2007. *Language Game of Responsible Agency and the Problem of Free Will: How can epistemic dualism be reconciled with ontological monism?* Philosophical Explorations, Vol. 10, No. 1.

---- 2008. *Between Naturalism and Religion*, trans. C. Cronin, Cambridge: Polity. [2005]

---- 2010. *An Awareness of What is Missing. Faith and Reason in a Postsecular Age*, trans.C. Cronin. Cambridge: Polity Press.

Taylor, C. 'The Motivation Behind a Procedural Ethics', in *Kant and Political Philosophy: The Contemporary Legacy*, Ronald Beiner and William James Booth ed., New Haven, Yale University Press, 1993, pp. 337-360.

--- 'Overcoming Epistemology' in *Philosophical Arguments*, Cambridge, MA: Harvard University Press, 1995, pp. 1-19.

--- 'The Validity of Transcendental Arguments', In *Philosophical Arguments*. Cambridge, MA: Harvard University Press, 1995. 20-33

--- 'Irreducibly Social Goods', In *Philosophical Arguments*. Cambridge, MA: Harvard University Press, 1995. 127-145.

--- 'Comparison, History, Truth', In *Philosophical Arguments*. Cambridge, MA: Harvard University Press, 1995. 146-165.

---- 'Explanation and Practical Reason', In *Philosophical Arguments*. Cambridge, MA: Harvard University Press, 1995. 34-60.

--- *The Ethics of Authenticity,* Cambridge: Harvard University Press, 1992.

--- *Sources of the Self: The Making of the Modern Identity*, Cambridge: Harvard University Press, 1989.

--- 'What is Human Agency?', In *Human Agency and Language: Philosophical Papers 1*. Cambridge: Cambridge University Press, 1985. 15-44.

--- *Philosophy and the Human Sciences: Philosophical Papers*, Vol 2. Cambridge: Cambridge University Press.

--- 'Foucault on Freedom and Truth', In *Philosophy and the Human Sciences: Philosophical Papers 2*. Cambridge: Cambridge University Press, 1985, pp.152-184.

--- ‘Le Fondamental Dans l'Histoire’, In *Charles Taylor Et l’interprétation De l’identité Moderne*, edited by Guy Laforest and Philippe de Lara. Paris;Sainte Foy;: Centre Culturel International de Cerisy-la-Salle, Cerf;Les Presses de l'Université Laval;, 1998. 35-51.

Secondary Literature:

Adorno, T. W. 1973. *Negative Dialectics*, trans. E. B. Ashton, New York: Seabury Press.

Agar, N. 2004. Liberal Eugenics: In Defense of Human Enhancement. Oxford: Blackwell Publishing.

Allen, A. 2003. “Foucault and Enlightenment: A Critical Reappraisal” in *Constellations* 10, no. 2: 180–98.

---2008. *The Politics of Our Selves: Power, Autonomy, and Gender in Contemporary Critical Theory*, New York: Columbia University.

---2010. ‘The Entanglement of Power and Validity’. *Foucault and Philosophy*, O’Leiry (ed). Cambridge: Blackwell.

---2011c. “Foucault and the Politics of Our Selves” in Koopman (ed.), “Foucault Across the Disciplines” special issue of *History of the Human Sciences* 24, no. 4, Oct.: 43–60.

Amiot. M. 1967**.** *Le relativisme culturaliste de Michel Foucault*, Les Temps Modernes, n. 248, pp. 1271-1298.

Arens, E. (ed.), *Habermas und die Theologie. Beiträge zur theologischen Rezeption, Diskussion und Kritik der Theorie kommunikativen Handelns* (Düsseldorf: Patmos, 2nd edn, 1989), trad.it. *Habermas e la teologia*, Queriniaina, Brescia, 1992.

Benhabib. S. 1992. *Situating the Self*. London: Routledge.

Bernauer. J. W. 1990. *Michel Foucault’s Force of Flight: Toward an Ethics for Thought*. London: Humanities Press.

Buchanan, A., D. Brock, N. Daniels, and D. Wikler. 2000. From Chance to Choice–Genetics and Justice. Cambridge, Cambridge University Press.

_____ 2011 Beyond Humanity?: The Ethics of Biomedical Enhancement**.** Oxford, Oxford University Press.

J. Butler, ‘Critique, Dissent, Disciplinarity’ in conception of critique in modern and contemporary philosophy, Houndmills, Palgrave Macmillian, 2012

R. Caillois [1958], *Man, Play and Games*, Free Press of Glencoe, New York, 1961.

Christiansen, K. 2009 The Silencing of Kierkegaard in Habermas' Critique of Genetic enhancement, Med Health Care and Philos, 12:147–156.

W. Connolly, 'Foucault and Otherness,' *Political Theory*, Vol. 13, No. 3 (Aug., 1985), pp. 365-376.

M. Cooke, 'The Limits of Learning: Habermas' Social Theory and Religion', in *European Journal of Philosophy*, 2014.

---------- 'Translating Truth' in *Philosophy and Social Criticism*, 2011, pp. 479-492.

Davidson, A "Archaeology, Genealogy, and Ethics," in Hoy, *Foucault A Critical Reader*, Blackwell, Oxford, Blackwell, 1986, pp. 221- 233.

Dews, P. (1989) 'The Return of Subjectivity in the late Foucault', Radical Philosophy 51: 37–41.

--------ed. 1999. *Habermas – A Critical Reader,* Oxford : Blackwell.

Dreyfus, H. and Rabinow, P, *Michel Foucault. Beyond Structuralism and Hermeneutics,* H. Dreyfus and P. Rabinow eds., Chicago. University of Chicago Press, 1983[2]

---------'What is Maturity? Habermas and Foucault on 'What is Enlightenment?' in *Foucault: A Critical Reader*, ed. D. Couzens Hoy, Oxford, Basil Blackwell, 1986.

Edgar, A. 2009. The hermeneutic challenge of genetic engineering: Habermas and the transhumanists. Medicine, Health Care and Philosophy.

R. Fillion, 'Foucault contra Taylor: Whose Sources? Which Self?' *Dialogue* (Fall 1995), pp. 663-74.

---------- 'Freedom, Truth, and Possibility in Foucault's Ethics', in Foucault Studies No. 3 (November, 2005), 50-64.

Fraser, N. (1989) Unruly Practices: Power, Discourse, and Gender in Contemporary Social Theory. Minneapolis: University of Minnesota Press.

Fukuyama, F. 2002. Our Posthuman Future—Consequences of the Biotechnology Revolution (trans: Rehg, W., Pensky, M., and Beister, H.) London: Profile Books.

Glover, J. 1984. *What Sort of People Should There Be?* Harmondsworth: Penguin.

Gordon, P. E. "The Place of the Sacred in the Absence of God: Charles Taylor's *A Secular Age*," *Journal of the History of Ideas* 69, no. 4 (2008), 647-673

G. Gutting, 'Foucault, Hegel and Philosophy', in *Foucault and Philosophy*, Wiley-Blackwell, Chichester, 2010.

Han-Pile, B. (2002) Foucault's Critical Project: Between the Transcendental and the Historical, trans. E. Pile. Stanford, CA: Stanford University Press.

---------'The Analytic of Finitude and the History of Subjectivity', in *The Cambridge Companion to Foucault*, Cambridge University Press, Cambridge, 2005.

----------'The "Death of Man": Foucault and Anti- Humanism' in *Foucault and Philosophy*, Wiley-Blackwell, Chichester, 2010.

Harris, J. 2007. *Enhancing evolution—the ethical case of making better people*. New Jersey: Princeton University Press.

Herissone-Kelly, Peter (2012) ***Habermas, Human Agency, and Human Genetic Enhancement: The Grown, the Made, and Responsibility for Actions.*** Cambridge Quarterly of Healthcare Ethics, 21 (02). pp. 200-210.

A. Honneth, *The Critique of Power: Reflective Stages in a Critical Social Theory*, Cambridge, MIT Press, 1991, chap. 4 and 6.

Horkheimer, M. 1947. *Eclipse of Reason*, New York: Oxford University Press.

Koopman, C. 2010b. "Historical Critique or Transcendental Critique in Foucault: Two Kantian Lineages" in *Foucault Studies,* no. 8, February: 100–21.

----------2013. *Genealogy as Critique*. Bloomington: Indiana University Press.

Larmore, C. "How Much Can We Stand?," *The New Republic* (April 9, 2008);

S. Le Bon, *Un positiviste désespéré: Michel Foucault*, Les Temps Modernes, n. 248, 1967, pp. 1299-319.

Lewens, T. (2012) *Human Nature: The Very Idea*, Philosophy and Technology 25 (4):459-474.

MacIntyre, A. (1994) "Critical Remarks on The Sources of the Self by Charles Taylor," Philosophy and Phenomenological Research, 54, pp. 187–190.

McCarthy, T., "The Critique of Impure Reason: Foucault and the Frankfurt School," in *Ideals and Illusions: On Reconstitution and Deconstruction in Contemporary Critical Theory*, Cambridge, MIT Press, 1991;

----------'Kantian Constructivism and Reconstructivism: Rawls and Habermas in Dialogue', *Ethics* 105 (1994), pp 44–63.

Mendieta E. ed. 2013. *Habermas and Religion*, Cambridge : Polity.

M. Nussbaum, "Our Pasts, *Ourselves", in Philosophical Interventions: Reviews 1986-2011,* Oxford University Press, New York,2012, p. 91.

T. O'Leary, *Foucault and the Art of Ethics*, Continuum, London, 2002.

Raz, **J.** 1999. *Engaging Reason*. Oxford: Oxford University Press.

Reiman. J. *1990*. *Justice and Modern Moral Philosophy*. New Haven: Yale University Press

Rorty, R. (1994) "Taylor on Truth." In Philosophy in an Age of Pluralism: The Philosophy of Charles Taylor in Question, ed. J. Tully. Cambridge: Cambridge University Press, pp. 20–33.

Rose, N. 2007. *The politics of life itself—biomedicine, power, and subjectivity in the twenty-first century.* Princeton: Princeton University Press.

Sandel, J.M. 2007. *The case against perfection—ethics in the age of genetic engineering*. Cambridge: The Belknap Press of Harvard University Press.

Q. Skinner, "Who Are "We"? Ambiguities of the Modern Self", in *Inquiry*, v. 34, (2), 1991,

Tully, J., ed. (1994*) Philosophy in an Age of Pluralism: The Philosophy of Charles Taylor in Question*. Cambridge: Cambridge University Press.

-------(1999)"To Think and Act Differently", in *Foucault contra Habermas*, Sage Publications, London.

Vanheeswijck, **G**. **"**Does History Matter? Charles Taylor on the Transcendental Validity of Social Imaginaries**."** History and Theory 54, no. 1 **(**2015**):** 69-85**.**

M. Walzer. 1980. "The Moral Standing of States", in *Philosophy and Public Affairs*, Vol. 9, No. 3, pp. 209-29.

--------1986. "The Politics of Michel Foucault," in David Hoy (ed.), *Foucault: A Critical Reader.* Oxford: Blackwell.

Wasserman, D. 2003. *My fair baby: What's wrong with parent's genetically enhancing their children? In Genetic prospects*—essays on biotechnology, ethics and public policy, ed. V. Gehring, s.273–s.294. London: W.W. Norton & Company.

Woodford, P. "Specters of the Nineteenth Century: Charles Taylor and the Problem of Historicism," *Journal of Religious Ethics* 40, no. 1 (2012), 171-192;

Zylinska, J. (2010) *Playing God, playing Adam: The politics and ethics of enhancement*, Journal of Bioethical Inquiry 7 (2):149-161