

# Conscience and Confession in Rousseau's Naturalistic Moral Psychology

*Naturalism Workshop*  
Columbia, South Carolina  
30-31 March, 2007

Wayne M. Martin  
University of Essex  
[wmartin@essex.ac.uk](mailto:wmartin@essex.ac.uk)

IN PLACE OF AN ABSTRACT: I here report on my work-in-progress addressing Rousseau's naturalistic account of human agency. In the first half of these notes I attempt to throw light on the distinctive character of Rousseau's philosophical naturalism. I compare Rousseau's naturalism both to that of his own contemporaries and to some of our own (§1), but argue that Rousseauian naturalism is better understood as a development of ancient forms of ethical naturalism, particularly as mediated by Seneca (§2). I then turn to consider how Rousseau's distinctive naturalistic commitments shape his treatment of the problem of self-consciousness, in particular with regard to the self-consciousness involved in action. I argue that Rousseau identifies two fundamental structures of self-consciousness essential to beings with natures like ours. The first is Rousseauian conscience, understood following the Stoics as a form of natural self-sentiment (§3); the second is associated with the distinctively human task of confession, understood as a form of self-judgment (§4).

## §1: Rousseau against the Naturalists

The first piece of our puzzle concerns the character of Rousseauian naturalism. Here I find it useful to contrast Rousseau both with his own contemporaries and with ours. For the forms of naturalism of concern in our own time I am using David Papineau's formulation in *Philosophical Naturalism*, together with a couple of influential ideas from Quine as taken up by various neo-Quineans. In the Introduction of the *Naturalism* book Papineau goes through the usual hand-wringing about the fact that naturalism is a broadly shared but poorly defined conviction. But unlike Barry Stroud, who once concluded that the term means nothing at all ("The Allure of Naturalism" APA Pacific Presidential Address, 1996), Papineau goes on to identify it with some fairly specific and substantial philosophical commitments. Two of these concern me here: the rejection of dualism in favor of some form of physicalism, and the insistence that philosophical inquiry is continuous with (but not foundational for) empirical natural science. To this I want

to add a third, which is really a corollary of the first two but deserves separate mention nonetheless: a kind of secularism, even atheism in philosophy. For the modern naturalist, the appeal to God, to what Hume called 'occult qualities,' or to what Patricia Churchland famously called 'spooky stuff' ("Can Neurobiology teach us anything about consciousness?" APA Pacific Presidential Address, 1993) has no place in respectable metaphysics or philosophy of mind.

There is also a fourth plank that one often finds in contemporary naturalism, though significantly Papineau is smart enough to avoid committing himself to it explicitly. As a canonical marker for the fourth plank we can take a famous passage from Quine's manifesto, "Epistemology Naturalized." Recall there Quine's final retort to Carnap's program: "But why all this creative reconstruction, all this make-believe? The stimulation of his sensory receptors is all the evidence anybody has to go on, ultimately, in arriving at his picture of the world. Why not just see how this construction really proceeds? Why not settle for psychology?" I don't want to take a stand on the exact commitment Quine himself meant to be undertaking with these notorious rhetorical questions. But part of the legacy of the Quinean stance for the subsequent naturalist tradition has been a privileging of a realistic, bio-mechanical investigation of the way in which embodied organisms actually process inputs to yield outputs. The exact nature of this privilege can vary substantially, of course. Minimally, a biologically accurate theory might act as a constraint on a theory of mind (that is, our philosophy of mind had better not be inconsistent with it). A more extreme position is that such a bio-physiological theory (or perhaps a set of them) simply amounts to the whole story there is to tell about mindedness.

This particular package of naturalistic commitments may sound distinctly contemporary, but it is in fact a form of naturalism with which Rousseau was intimately familiar. In particular, the more radical figures in the so-called Encyclopaedist movement -- here I have in mind in particular d'Holbach, La Mettrie and Helvétius (the 'French Materialists') -- closely adhered to these four commitments of contemporary naturalism. D'Holbach's substantial two-volume *System of Nature* (1770) is probably the crowning achievement of this movement. It was published only after Rousseau's major writings were complete, but Rousseau was undoubtedly familiar with much of d'Holbach's approach and program. For much of his life he was on quite intimate (though not always friendly!) terms with Holbach and frequented his literary and philosophical circle in Paris.

The first section of *The System of Nature* is simply called "Nature and Her Laws"; in it d'Holbach lays out his agenda. Allow me to quote the whole first paragraph:

Man has always deceived himself when he abandoned experience to follow imaginary systems.--He is the work of nature.--He exists in Nature.--He is submitted to the laws of Nature.--He cannot deliver himself from them:--cannot step beyond them even in thought. It is in vain his mind would spring forward beyond the visible world: direful and imperious necessity ever compels his return--being formed by Nature, he is circumscribed by her laws; there exists nothing beyond the great whole of which he forms a part, of which he experiences the influence. The beings his fancy pictures as above nature, or distinguished from her, are always chimeras formed after that which he has already seen, but of which it is utterly impossible he should ever form any finished

idea, either as to the place they occupy, or their manner of acting--for him there is not, there can be nothing out of that Nature which includes all beings.

Should we call this d'Holbach's naturalism? Before providing analysis, let me add in a couple of quotes from some others in d'Holbach's circle – the group Rousseau refers to in *The Confessions* as *la coterie d'Holbach*: the Holbach clique. In an elegant little book called *Machine Man* (1747), d'Holbach's ally, La Mettrie concludes: "Let us then conclude boldly that man is a machine and that there is in the whole universe only one diversely modified substance." What is that substance? Neither Mettrie nor Holbach seem to have thought of it in Spinozist terms. For them, the one substance is simply 'matter.' Here is Holbach again, in that first section of *The System of Nature*:

The universe, that vast assemblage of every thing that exists, presents only matter and motion: the whole offers to our contemplation, nothing but an immense, an uninterrupted succession of causes and effects . . . .

There is considerable overlap between these ideas and the commitments we find at work in more recent varieties of naturalism. D'Holbach and La Mettrie share the modern rejection of dualism in the philosophy of mind; materialism is their 18th century equivalent of physicalism; they reject any appeal to supernatural forces in metaphysics; they firmly hold that philosophy should become continuous with (indeed should simply become) a form of natural science.<sup>1</sup> And in all this there is the now-familiar but then-outrageous secularism; indeed d'Holbach and Mettrie (along with their contemporary and follower, the Marquis de Sade) may well be the first self-professed atheists in the philosophical canon.

What I distinguished above as the methodologically prescriptive fourth plank of naturalism -- the privileging of bio-mechanical explanation in philosophical psychology and epistemology -- was also a major part of the d'Holbach package of commitments. If we wish to understand the capacities of man, he held, or to improve the human lot, then we must seek out the distinctive material causes that produce the distinctive material effects associated with this distinctive configuration of matter known as 'Man'. The ultimate concerns of the French Materialists lay firmly in the area of social reform, but they were convinced that any attempts at social reform were simply so many more causal interventions in a vast causal mechanism. If they were to be intelligent and effective interventions they must be grounded in an understanding of the intricate mechanism in which they sought to intervene. Here is one last quote from D'Holbach: "All the steps taken by man to regulate his existence, ought only to be considered as a long succession of causes and effects, which are nothing more than the development of the first impulse given him by nature."

So if Rousseau knew a close cousin of modern naturalism, what did he make of it? Well I think he hated it. Certainly he had a strong dislike for Holbach himself. Rousseau describes his reaction upon first meeting Holbach as one of "natural repugnance," and *The Confessions* are full of nasty personal anecdotes about him. In many ways this is odd. By most accounts d'Holbach was an exceptionally generous man

---

<sup>1</sup> "I imagined that morality ought to be treated like all the other sciences, and founded on experiment, as well as natural philosophy." Helvétius, Preface to *De L'Esprit* (1758).

and a kind host. He used his considerable personal wealth to support the Encyclopedia, and his home was the site of many lively and stimulating parties and discussions, in which Rousseau was regularly included. Rousseau himself admits that Holbach went out of his way to welcome him, well before his rise to literary prominence. But Rousseau seems not to have returned the kindness. He reports that d'Holbach sensed this hostility and finally confronted him about it, asking: "Why do you flee from me?" Rousseau's pithy answer: "You are too rich."

But what did Rousseau make of d'Holbach's philosophical project? The question is not entirely straightforward to answer, but I think it is fair to say that he rejected all the central naturalistic commitments of French materialism. Certainly his published writings include explicit rejections of a materialist ontology in favor of a dualistic one. Rousseau: "as for me, whatever Locke says about it, I need only know that matter is extended and divisible in order to be sure it cannot think" (*Émile*, p 279 in Bloom's translation). One must exercise some care here, however; the passage just cited comes from the Creed of the Savoyard Vicar, so we must remember that the first-person pronouns in the foregoing sentence refer to one of Rousseau's characters. But note that in the context of the Second Discourse, for instance, which is presented in the author's own voice (albeit anonymously), Rousseau is perfectly happy to enumerate the "metaphysical" and "moral" properties of original man in addition to his physical traits. D'Holbach explicitly rejected such distinctions as inconsistent with his materialist approach.<sup>2</sup>

So it seems that Rousseau is no physicalist; neither is he an atheist. He says in the letter to Beaumont that he is a Christian, although he is careful to describe his Christianity as the Christianity of Christ rather than that of the church. The first principle in the Vicar's Creed postulates "a will that moves the universe and animates nature" (*Émile*, 273). The society of the Social Contract mandates religious observance. Rousseau seems to have viewed the atheism of his contemporary naturalists as both a mistake in metaphysical theory and as a real danger in moral and political practice; he saw their belittling dismissal of religious consciousness and religious observance as threatening a real and irreplaceable loss. [Is this true? I need to add some textual evidence in support of the hunches expressed in this paragraph. In the literature Rousseau is variously described as a Deist or even as a Pantheist. In his own day these views were regularly equated with atheism, but they are surely not to be equated with the avowed atheism of his contemporaries or ours.]

A third point of divergence from the naturalistic program is to be found in Rousseau's methodological pluralism. Both the 18<sup>th</sup> and the 20<sup>th</sup> century naturalisms we have considered are methodologically prescriptive, advocating some privileged method of addressing and resolving philosophical questions. The exact prescription varies from one naturalist to another: D'Holbach preferred

---

<sup>2</sup> D'Holbach: "The distinction which has been so often made between the *physical* and the *moral* being, is evidently an abuse of terms. Man is a being purely physical: the moral man is nothing more than this physical being considered under a certain point of view; that is to say, with relation to some of his modes of action, arising out of his individual organization."

material causal explanation; his modern counterparts variously insist on behaviorism or neurophilosophy or heterophenomenology. (These are, of course, the methodological prescriptions of Ryle, Churchland, and the recent Dennett respectively.) Nothing could be further from Rousseau's philosophical and literary sensibilities than this sort of methodological monism. Where Papineau calls for a philosophy that is 'continuous with empirical science,' Rousseau famously begins his *Discourse on Inequality* by insisting that "we must set aside all the facts, which have no bearing on the question," and specifically enjoining his readers to "leave aside all the scientific books which teach us only to see men as they have made themselves." (At least in some of his moods, it seems, Rousseau is perfectly happy with philosophical 'make-believe'!) He certainly draws on the results of natural science (botany is his favorite science, but he was also deeply influenced by what would later be called ethology: the study of animal behaviour), but he is just as likely to address a philosophical problem by constructing a wildly counter-factual state-of-nature narrative or a Reverie or an auto-biography or a novel. For Rousseau there simply is no one single privileged method of philosophical investigation or exposition.

In exploring Rousseau's antipathy to the naturalism of his materialist contemporaries, there is one work in particular that looms especially large. Helvétius' book *De L'Esprit* ('On Spirit' or 'On the Mind') was published just in the period of Rousseau's intensive work in bringing *Émile* to completion. (*De L'Esprit* appeared in the summer of 1758; Rousseau had completed a draft of *Émile* -- the so-called *Manuscrit Favre* -- by 1759.) *De L'Esprit* is not now a well-known book. (The last English translation, as far as I can tell, was the one published in London in 1759.<sup>3</sup>) But in its day it was notorious: "condemned by the Sorbonne, the Pope and the Parliament of Paris and burnt by the public executioner", according to one modern description; "atheistic, materialistic, sacrilegious, immoral and subversive," according to another.<sup>4</sup> When Rousseau read it he was clearly provoked. He wrote extensive notes in the margins of his personal copy, and made plans to publish a rebuttal. But he soon realized that his principled objections would simply be lost in the heat of the *scandale*, and so decided instead to recast some key passages in *Emile* to present his rebuttal of Helvétius' materialist theory of mind.<sup>5</sup>

The controversy between Rousseau and Helvétius merits close scrutiny – closer scrutiny than I shall be able to provide here. For present purposes I confine my comments to a few remarks about the fundamental issue of dispute. Rousseau's most insistent criticisms of *De L'Esprit* were focused on one of its leading ideas: the claim, as Helvétius put it, that "to judge is nothing but to sense".<sup>6</sup> For the Materialists

---

<sup>3</sup> *De L'Esprit or Essays on the Mind and its Several Faculties*. The translation was printed in London in 1759, identifying no publisher or translator, but listing various booksellers who were licensed to sell it (four in London and one each in Oxford and Cambridge respectively). That first early translation has been reset various times but never superseded. A photo-mechanical reproduction of an 1810 edition was recently published by Kessinger Reprints, 2004.

<sup>4</sup> Smith 1964: 332.

<sup>5</sup> Explicitly the text of *De L'Esprit* remains neutral on the issue of materialism. But this was seen both by its contemporary readers and by modern scholars as an insincere gesture intended to get the book approved by the censors. For a discussion of the extraordinary circumstances of publication, see Smith 1965: *Helvétius: A Study in Persecution* (Oxford: Clarendon).

<sup>6</sup> "Juger n'est jamais que sentir." Helvétius 1758: Essay I, Ch. 1.

this claim was crucial to collapsing the distance between what Descartes had called “*la bête machine*” and the higher psychological capacities of human beings. If sensation could be accounted for in materialist terms, and if judgment was nothing more than a higher order variety of sensation, then all human psychology could be seen as the product of so many configurations of matter. Accordingly, Helvétius sets about the task of accounting for judgment as an essentially passive psychological phenomenon – the registration of similarities and difference. For Rousseau, however, such an approach effectively effaced what he took to be the essential character of judgment. Again and again in *Émile* we find Rousseau returning to the contrast between sensation and judgment, in what was clearly intended as a direct reply to Helvétius. “Our sensations are purely passive, while all our perception or ideas are born out of an active principle which judges” [*Émile*, 107]. “In sensation, judgment is purely passive. ... In perception or idea, judgment is active” [*Émile*, 203].

The two passages just cited come from outside the Creed, but the fullest statement of the point is found within it:

Judging and sensing are not the same thing. By sensation, objects are presented to me separated, isolated, such as they are in nature. By comparison I move them, I transport them, and, so to speak, I superimpose them on one another in order to pronounce on their difference or their likeness and generally on all their relations. According to me [i.e., the Vicar], the distinctive faculty of the active or intelligent being is to be able to give a sense to the word *is*. I seek in vain in the purely sensitive being for this intelligent force which superimposes and which then pronounces; I am not able to see it in its nature. This passive being will sense each object separately, or it will even sense the total object formed by the two; but having no force to bend them back on one another, it will never compare them, it will not judge them. [*Émile*, 270-271.]

In short, what Rousseau found most objectionable in the naturalism of the French Materialists was what he saw as its misrepresentation of judgment, and hence its inability to account for all manner of other distinctively human capacities – notably thought, reason and action. Even the most minimal perceptual judgment, Rousseau insists, involves psychological operations that go beyond the mere registration of sensory content. It involves first the ability to undertake comparisons among sensations; it then involves the capacity to make an objective claim -- to exercise our understanding of the word ‘is.’ These two capacities – what Rousseau here describes as “the intelligent force which superimposes and which then pronounces” [*cette force intelligente qui superpose et puis qui prononce*] – are powers that can be exercised on sensory contents, but which are not themselves sensory capacities. As Rousseau writes in the margin of his personal copy of *De L’Esprit*: “the comparison of yellow and red is not the sensation of yellow nor that of red.”<sup>7</sup>

But underlying these two extrasensory components of judgment Rousseau finds a more fundamental capacity for self-determination, and it is here that we find his deepest divergence from the naturalism of the Materialists. For what ultimately prevents the Materialists from accounting for judgment, according to

---

<sup>7</sup> Rousseau, “Notes on Helvétius’s *On the Mind*”; *Collected Writings of Rousseau*, Vol. 12 (2007), p. 205.

Rousseau, is what Holbach himself called the commitment to “direful and imperious necessity,” and the consequent treatment of human beings as essentially passive or receptive nodes in a mechanical causal system. Like Descartes and Kant, Rousseau held that such an account could never account for the distinctively human capacity for judgment, which involves an irreducibly active, autonomous principle. Here is one more passage from the Vicar’s Creed: “And what is the cause which determines his will? It is his judgment. And what is the cause which determines his judgment? It is his intelligent faculty, it is his power of judging: the determining cause is in himself. Beyond this I understand nothing more.” (*Émile*, 280)<sup>8</sup>. In the Creed this is exactly the line of thought that is used to motivate the third, explicitly anti-materialist article of faith: “Man is therefore free in his actions and as such is animated by an immaterial substance. This is my third article of faith” (*Émile* 281).

It is not my purpose in this discussion to defend this line of argument. Indeed in the last analysis we have to conclude that Rousseau was not all that careful as a metaphysician. Note, for instance, that his rejection of materialism in the Creed is built on the basis of some very bad physics, even by 18<sup>th</sup> century standards. (“It is the essence of matter to remain at rest.”) The theology of the Creed is informed by a version of the teleological argument that Hume would shortly dismantle; much of his positive metaphysical argumentation about the self is paralogistic. And the idea that substance dualism somehow makes freedom intelligible is one that, well ... I don’t really understand and wouldn’t want to endorse. But I’ll settle here for two interpretative points that I shall develop in the remainder of the discussion: first that Rousseau rejects Papineau/Holbach naturalism, and secondly that he holds that what it specifically leaves unexplained is the capacity for judgment, which requires a conception of human beings as essentially active and free.

Before leaving this topic, it is worth adding one further note about Rousseau’s reaction to Helvétius. As we have seen, the main thrust of Rousseau’s objection pertains to the account Helvétius provides of judgment. But there are also traces of a second line of criticism, this one pertaining to memory. The account of memory in *De L’Esprit* is exactly analogous to the treatment of judgment. Here again the materialists’ strategy was to reduce memory to a form of sensation. Here again Rousseau recoils. Where Helvétius writes “to remember, as I am going to prove, is properly only to feel”, Rousseau writes in the margin: “I do not yet know how he is going to prove that; but I know very well that to feel the present object and to feel the absent object are two different operations whose difference very much deserves to be examined.” When Helvétius goes on to conclude his proof “it is then evident that to remember is to feel”, Rousseau writes in the margin: “there is this difference that the memory produces a similar sensation and not the sentiment, and this other difference also that the cause is not the same.” And in yet a third marginal comment Rousseau writes that on Helvétius’ view, “it would be impossible to distinguish the memory of the sensation from the sensation.” The details of Rousseau’s line of attack here require further study, but

---

<sup>8</sup> See also *Émile* 275: “to act, to compare and to choose are operations of an active and thinking being.” [check]

we can at least say this: in memory and in judgment Rousseau found two fundamental psychological phenomena that he held to be distorted when forced into the framework required by the naturalistic materialism of his day.

## §2: Neo-Stoic Naturalism

So how *should* we conceive of Rousseau's naturalism, given his antipathy towards the Papineau/d'Holbach variety? I shall not here attempt a complete answer to this rather complex question, but I would like at least to suggest a strategy for thinking about it. The suggestion in effect comes from Rousseau himself. As the motto for *Émile*, Rousseau uses a passage from Seneca's essay *On Anger*: "We are sick with evils that can be cured; and nature, having brought us forth sound, itself helps us if we wish to be improved." Seneca was of course one of the leading Stoic writers of the Roman world; he was also a considerable influence on Rousseau. I propose that in thinking about Rousseau's naturalism we will do well to set aside the forms of naturalism predominant among his own contemporaries and among our own, and think instead of Rousseau's project in the context of an ancient form of naturalism which Rousseau set out to appropriate and modernize. It is striking that amid the considerable recent scholarship on Rousseau – much of it excellent – there has been a good deal more attention paid to his influence than to his influences. Rousseau's legacy in the thought of Kant, Goethe, Hegel, Marx, Freud, ... is now well-documented and much-studied. But his distinctive appropriation of ancient sources, particularly of ancient Roman sources, is much less well understood. In thinking about Rousseau's naturalism, I submit, these ancient sources are of considerable importance.

In developing this suggestion it will be useful to start by introducing a generic naturalistic position that was, I believe, very widely shared among the ancients. Let's call it *Ancient Consensus Naturalism*. Ancient Consensus Naturalism is a thesis about living well. In particular it is the thesis that one cannot live well if one lives contrary to one's nature. This is a conviction that is rarely (if ever?) argued for directly among the ancients, but it is one that operates as an assumption in a great deal of ancient moral psychology, which is characteristically concerned with figuring out just what forms of excellence one's nature requires. Notice first of all that Ancient Consensus Naturalism is a form of what I will call ethical or normative naturalism. Unlike the modern naturalisms we have considered so far, it is not in the first instance a thesis in metaphysics, epistemology, or the philosophy of mind. Secondly, notice that as formulated, Ancient Consensus Naturalism is a *negative* thesis – indeed doubly so. One *cannot* live well by living *against* one's nature. But the two negatives here do not add up to a positive. Ancient Consensus Naturalism does not guarantee that living in accord with one's nature will yield *eudaimonia*. In fact I think that most ancient moralists (here I am thinking not only of the philosophers but also of the ancient poets and playwrights) were agreed that there is no such guarantee. After all, a lack of external goods or a spate of bad luck or

some unwelcome intervention from the gods might keep me from *eudaimonia*, even if I am living in accord with my nature. Notice thirdly that although it is primarily an ethical doctrine, Ancient Consensus Naturalism is committed to a substantial and controversial metaphysical thesis: namely the claim that there is in fact such a thing as human nature, and moreover that it is defined teleologically. If physicalism is the metaphysical bedrock of 20<sup>th</sup> century naturalism and materialism was the bedrock of its 18<sup>th</sup> century predecessors, it is metaphysical essentialism upon which ancient naturalism was erected.

Ancient Consensus Naturalism was, I believe, a default position among ancient moralists. But there was a stronger form of ethical naturalism found in the ancient world that was anything but a consensus view. Here I have in mind the notorious Stoic thesis that the life lived in agreement with nature simply is the highest good and final end of human life. According to one ancient anthologist, this was the position progressively articulated over the first generations of Stoic teachings:

Zeno represented the end as living in agreement [*ho-mo-lo-GOU-menos zen*]. ... His successors expressed this in a more expanded form, 'living in agreement **with nature**,' [*ho-mo-lo-GOU-menos **tei phusei zen***] -- since they took Zeno's statement to be an incomplete predicate. [It was] Cleanthes, Zeno's first successor, [who] added 'with nature' and represented the end as follows: 'the end is living in agreement with nature.'<sup>9</sup>

The accuracy of Stobaeus' history has been contested, but need not concern us here. For our purposes what matters is that by the time of the leading Roman Stoics, the principle *sequi naturam* ('follow nature') had come to be treated as an overarching ethical maxim and a core commitment of Stoic moral philosophy. It forms the kernel of what I shall call Stoic Naturalism.

Stoic Naturalism is Ancient Consensus Naturalism *plus*. It is committed to the negative thesis so common among ancient moralists, and it is accordingly committed to the controversial metaphysical thesis that there is indeed a teleological human essence. But Stoic Naturalism goes further. In its strongest form it holds that living in accordance with nature is *both necessary and sufficient* for happiness. Since the Stoics also held that living in accordance with nature *is* virtue, they notoriously adhered to the intensely controversial and much-ridiculed view that virtue is sufficient for happiness. The good man, it seems, will be happy even on the rack of torture.

The thesis I would like to propose here is that Rousseau's naturalism should be understood as a distinctive development and modification of Stoic Ethical Naturalism. This is not to say that Rousseau was himself a Stoic; indeed most of the explicit mentions of Stoicism in Rousseau's corpus are decidedly negative.<sup>10</sup> Nonetheless Rousseau clearly borrowed very heavily from the Stoics, and in particular from Seneca. The argument of the First Discourse takes both its thesis and much of its argumentative method from Seneca's 88<sup>th</sup> Letter ("On Liberal and Vocational Studies"), from which Rousseau also adapted his famous *bon mot* about private property. Parts of the Second Discourse are borrowed from Seneca's 90<sup>th</sup>

---

<sup>9</sup> Stobaeus, *Eclogues* 2.75,11-2.76,8; translation from Sedley and Long, *The Hellenistic Philosophers*, (Cambridge: Cambridge University Press, 1987), I, 394.

<sup>10</sup> In *Émile*, for instance, we find him questioning the notoriously restrictive Stoic account of the constituents of the good: "Why would I want to be Cato, who disembowels himself, rather than Caesar triumphant?" (*Émile*, 287).

Letter (“On the Part Played by Philosophy in the Progress of Man”). Indeed by modern standards Rousseau’s use of Seneca might well be treated as plagiarism; as one contemporary critic noted, the text of the First Discourse uses a verbatim quote from Seneca without attribution!<sup>11</sup> But it is not my purpose here to convict Rousseau of plagiarism, nor indeed to undertake a general survey of Seneca’s influence on his works. I am concerned rather with the ways in which variants on ancient forms of naturalism, and in particular Stoic naturalism, reappear at the center of Rousseau’s project.

So where does Rousseau stand with regard to these ancient forms of naturalism? Consider in turn the two doctrines we have distinguished. Although the evidence is not entirely unambiguous, I believe that there is strong reason to suppose that Rousseau endorses Ancient Consensus Naturalism. For our purposes here I rely on one memorable and representative passage from the fifth of the Letters to Sophie:

If moral goodness is in conformity to our nature, man cannot be healthy or well constituted unless he is good. If it is not and man is naturally wicked, he cannot cease being so without corrupting himself. ... [A] humane man would be as deprived an animal as a wolf with pity and only virtue would leave us with remorse. Would you believe that there could be an easier question in the world to resolve?<sup>12</sup>

Notice first that the disjunction Rousseau here proposes relies implicitly on the controversial essentialist assumption that we have seen to lie at the core of ancient normative naturalisms. Like the ancients, Rousseau seems to assume that there is such a thing as human nature, and that it is to be understood as a teleological structure – a norm that defines a standard of conformity and divergence. Moreover the very language that Rousseau uses here is suggestive of the legacy of ancient moral theory. Notice in particular the central role for the idea of being “healthy and well-constituted”<sup>13</sup> – as effective a translation of the ancient notion of *eudaimonia* as one is likely to find. If we also assume that Rousseau means to endorse the first of the two disjuncts proposed here, then we find in Rousseau’s moral philosophy a close variant of the ancient consensus: man cannot be *eudaimon* unless he is good, and such goodness involves conformity to one’s nature.<sup>14</sup>

---

<sup>11</sup> “Since learned men began to appear among us ... good people have slipped away.” The original source is Seneca’s 95<sup>th</sup> Letter (“On the Usefulness of Basic Principles.”) The source was identified by Canon Joseph Gautier in his public rebuttal of the First Discourse at a meeting of the Royal Society of Nancy, subsequently published in the October 1751 issue of the *Mercure de France*. See *The Collected Writings of Rousseau*, II, 81.

<sup>12</sup> *Moral Letters V*, *Collected Writings of Rousseau* 12: 193. For a very similar passage see *Émile*, 287.

<sup>13</sup> *sain ... bien constitué*

<sup>14</sup> As is usual with Rousseau, there is a lot of hidden complexity in this passage. We must remember first that Rousseau was famous for holding that animals (including wolves, presumably) *do show pity*. Moreover, we must keep in mind that the *Moral Letters* were written in an attempt to make amends to a married woman who had rebuffed Rousseau’s romantic advances. This is, to say the least, a rather specialized context for the conduct of moral psychological investigations, and it puts in quite a special light the question as to whether being virtuous might leave one with remorse. I shall not speculate further about the hidden meanings of the *Moral Letters*; suffice to acknowledge that I do not take the textual evidence provided here as decisive evidence for Rousseau’s endorsement of Ancient Consensus Naturalism. I do think, however, that this text and others like it form a strong *prima facie* case. Rousseau’s account of his complicated relation with Sophie (Elisabeth-Sophie-Françoise Lalive de Bellegarde, Comtesse d’Houdetot) is found in *Confessions* IX and X. The Sophie addressed in the *Moral Letters* is of course not to be confused with the character of the same name in *Émile*.

What about the stronger Stoic thesis? I think it would be a mistake to ascribe to Rousseau the doctrine of Stoic Naturalism in its canonical form. Certainly there is nothing to suggest that Rousseau endorses the most extreme consequences of the Stoic sufficiency thesis. To put the point in the terms of the ancient debate: Rousseau seems to recognize both genuine external goods and the ethical significance of luck. In one passage from the correspondence he specifically criticizes what he calls the Stoic confusion of virtue and happiness.<sup>15</sup> But if Rousseau's naturalistic position should not simply be identified with Stoic Naturalism, it is, I submit, deeply indebted to it and is best understood with reference to it. One measure of the Stoic legacy in Rousseau's thought can be found in his adoption of the maxim, *sequi naturam*. ("Do you wish always to be well guided? Then always follow nature's indications."<sup>16</sup>) But to get a finer sense for Rousseau's Stoic naturalism we will do well to take our orientation from the motto he adopts from Seneca.

"We are sick with evils that can be cured; and nature, having brought us forth sound, itself helps us if we wish to be improved." It is striking just how much of Rousseau's distinctive philosophical position is summed up here in a single sentence. Both Seneca and Rousseau reject the orthodox Christian thesis that human beings are naturally and originally corrupt. On the contrary, they insist on the original goodness of man, who has been "brought forth sound." Both also agree in finding human beings to be quite thoroughly alienated from this natural goodness – corrupted, sick, evil. But what is most important for our purposes is the distinctive way in which both look to nature for a kind of moral therapy: a natural cure for our moral ailments. This confidence – that nature not only made us good but also provides us with the means for the recovery of our moral health – is part of the distinctive Stoic legacy in Rousseau's naturalism. For those (like us) who suffer from moral corruption and alienation, following nature cannot simply be a matter of following our corrupted and denatured instincts. Nevertheless, both Seneca and Rousseau are confident that we can find in our nature resources both for understanding our distinctive illness and for recovering our natural health. What Rousseau takes from Seneca's Stoicism, I propose, is far more than a few plagiarized passages; it is this schema for his whole moral psychological project.

But Rousseau does not simply take over this Stoic schema in order to reiterate Stoic doctrines. His challenge, as I would like to frame it, is to apply this ancient naturalistic schema in a specifically *modern* context. The issues here are complex, but for our purposes I would like to emphasize two related points. Already in the *Second Discourse* Rousseau had emphasized that human nature is distinctive in what he calls its 'perfectibility', or what we might now call its 'plasticity.' Unlike (most?) other animals, which simply live out their lives in accordance with a plan fixed by their nature, human beings exhibit quite different modes of existence depending on the context in which they find themselves. Being true to one's

---

<sup>15</sup> He warns his correspondent not to "exaggerate matters beyond the truth, nor to confound, as the Stoics did, happiness with virtue." (Rousseau to M. D'Offreville, 4 October, 1761). [But see the exchange over just this point at *Émile*, 282.]

<sup>16</sup> *Émile*, 363. See also *Moral Letters* V, CW 12, 194: "Let us obey nature, we shall know with what sweetness it approves what it has commanded ... ."

nature thus cannot have the same significance for a human being as it does, say, for a koala or a salmon, since human nature rather dramatically underdetermines the mode of existence of particular human beings. For Rousseau this underdetermination runs very deep, encompassing not only modes of dwelling and dress, language and diet, social and political and economic structure, etc., but also the facts about human moral and emotional life, the sentiments to which we are subject, the motives which incline us to act. This is one aspect of a broader problem that is arguably the central theme of all of Rousseau's writings: the problem of freedom. As we have already had occasion to note, Rousseau takes freedom to be a central and distinctive part of the human endowment. Once again this creates a paradox for applying the Stoic ethical schema; what does it mean to follow nature if one's nature is to be free?

With this we have arrived at some very large questions – too large, indeed, to be tackled responsibly here. Both of these problems have deep roots in the Stoic tradition, which arguably lacked an adequate solution to either of them. Both also figured centrally as part of Rousseau's legacy in the Idealist tradition in Germany. Even within the compass of Rousseau's thought, the attempt to answer this challenge implicates important themes in his political and educational program and in his philosophy of history as well as in his moral psychology and ethics. I shall not attempt here to enter this thicket. Instead I propose to use the remainder of these remarks to address a narrower set of issues that bear quite directly on these broader problems. In particular, I want to consider how Rousseau deploys the schema provided by his neo-Stoic naturalism in accounting for the role of self-consciousness is human agency.

### §3 Conscience as Natural Self-Sentiment

For Rousseau, I want to argue, human self-consciousness comes in two fundamentally different varieties, each of which has an essential role to play in the life of a well-functioning natural human agent. The first of these is conscience, understood in Rousseau's idiosyncratic manner. Conscience is an absolutely central notion in Rousseau's ethics, and it is a theme which drives him to some of his most extravagant rhetorical heights.

Conscience, Conscience! Divine instinct, immortal and celestial voice, certain guide of a being that is ignorant and limited but intelligent and free; infallible judge of good and bad which makes man like unto God; it is you who makes the excellence of his nature and the morality of his actions (*Émile*, 290).

Both in the portion of the Vicar's creed dealing with ethics, and in the earlier *Moral Letters*, the admonition to follow one's conscience figures as a central moral principle: "Conscience never deceives; it is man's true guide" (*Émile*, 286). This is of course a controversial claim, and much of the argumentative work in Rousseau's explicitly ethical writings is devoted to an elaborate defense of it. In particular, Rousseau sets out to defend the authority of conscience from those skeptics who would find in it nothing but an

internalized voice of some authority figure (parent, priest, teacher, ...) dogmatically reinforcing the ultimately arbitrary mores of one's local community.

But at this word [*conscience*] I hear the clamor of those who are allegedly wise rising on all sides: errors of childhood, prejudices of education, they all cry in a chorus. (*Émile*, 289)

Rousseau associates this line of objection particularly with Montaigne, but it could just as well be attributed to an ancient skeptic or to a post-modern moral relativist.

It is in his attempt to answer this objection that Rousseau undertakes his controversial genealogy of conscience. In particular, he sets out to show that at its core, conscience is not a socially constructed and culturally variable guide; it is a natural and universal moral endowment, an apt and reliable guide to conduct, and indeed a form of normativity that humans share in certain fundamental respects with the animals. In short, he seeks to show that the maxim to *follow one's conscience* is a specification of the more general principle to *follow nature*: "he who follows conscience obeys nature and does not fear being led astray" (*Émile*, 286).

I shall not here attempt to reconstruct or defend the details of Rousseau's genealogy of conscience, which will likely strike modern readers as rather jejune. What matters for my purposes is not so much the question of whether Rousseau is right in his extravagant claims about the reliability of conscience; I am interested rather in his distinctive analysis of its phenomenological character and content. Here I find it useful to think of Rousseau's position as working against the predominant metaphor of conscience as a distinctive inner voice. The metaphor of conscience as voice is extremely pervasive – from the Biblical description of "the still small voice of God" to the representations of conscience in modern popular culture. In its most hackneyed version, the metaphor would have us think of conscience as the voice of some kind of inner double or homunculus, uttering either general moral truths ("Always tell the truth") or moral guidance about particular courses of action, whether prospectively ('Don't do that') or retrospectively ('You shouldn't have done that.'). As we have already seen, Rousseau himself makes use of this dominant metaphor, but in the end his account of conscience serves to undermine it. The crucial point here is that to treat conscience as a voice suggests that its medium of representation is linguistic and conceptual – in Rousseau's language: a medium of ideas. Among other things, this is subtly to privilege the skeptic's position regarding the origin of conscience. If conscience speaks a language it is natural to ask whose language it speaks; the answer will lead us back to some purportedly authoritative Other. Against all this, Rousseau insists that the medium of representation in conscience is neither language nor ideas but feeling or sentiment. "The acts of conscience are not judgments but sentiments ... . To exist, for us, is to sense; our sensibility is incontestably anterior to our intelligence" (*Émile*, 290).

But if conscience is a sentiment, then what is it a sentiment of? Here again Rousseau's account carries a distinctive inflection. We commonly think of conscience as primarily a medium for the representation of the demands of duty or right action. Once again this is an account of the deliverances of

conscience that fits well with the treatment of conscience as the internalized voice of an authority figure. After all, what those authority figures were always telling us is what (not) to do. But while Rousseau certainly accepts that conscience provides an account of the demands of moral action, he holds that it does so by way of a distinctive form of *self*-representation. Consider for instance the continuation of the passage just cited:

The acts of conscience are not judgments but sentiments. Although all our ideas come to us from the outside, the sentiments evaluating them are within us, and it is by them alone that we know the compatibility or incompatibility between us and the things we ought to seek or flee. ... Whatever the cause of our nature, it has given us sentiments suitable to our nature. (*Émile*, 290; see also CW 12, 196)

Notice the doubly reflective structure of conscience on this account. In attending to the sentiments of conscience I am attending to a sentiment whose source, as Rousseau has it, is ‘within me’ – in contrast to sensations and perceptions whose source is external. But these sentiments themselves refer back to me in a distinctive way, insofar as the distinctive content of these sentiments concerns the proper normative fit between me and the things and my environment. The sentimental promptings of my conscience are thus grounded, according to Rousseau, in an implicit comprehension of the kind of being I am and of my fit with the things around me that either threaten or benefit me. Rousseauian conscience is thus a form of non-discursive self-comprehension.

On both of these points, Rousseau is again effectively following the lead of the Stoics. I have argued for this thesis at length elsewhere and shall not repeat the details of my analysis here. Allow me simply to summarize a few crucial points. For the Stoics, one of the fundamental concepts of ethics and psychology is *Oikeiosis* (variously translated as *appropriation*, *conciliation*, *affinity*). *Oikeiosis* is a natural endowment of all organisms in virtue of which they are naturally and innately attracted to what benefits them and naturally repelled from what threatens them. Crucially for our purposes, the Stoics held that *Oikeiosis* either is or involves a form of self-consciousness – what Chrysippus called *syneidesin*, Hierocles called *synaesthesia*, and Seneca called *sensus sui*. The object of this self-consciousness is what the Stoics called the organism’s “constitution” (*sustasis*, *constitutio*). An organism’s constitution is its governing principle; in particular it is the principle of mereological unity which governs how the various spatially and temporally distributed parts of an organism function together to preserve and reproduce the whole. According to the Stoics, every animal is created by nature with an implicit, inarticulate, practical and normative comprehension (Seneca calls it a *sensus*) of its own constitution; it is in virtue of this implicit self-understanding that the organism is capable of using its limbs appropriately and guiding itself through its environment successfully. Since every different animal species has a different form of mereological

unity, it follows that each has a different constitution or governing principle, and hence that this self-consciousness takes a different form in different kinds of animals.<sup>17</sup>

On all these points we can see Rousseau's theory of conscience as occupying the space staked out by the Stoic position. First, conscience is to be understood not as a linguistic medium or as a stock of moral ideas or principles but as a stock of feelings or sentimental promptings to act or refrain from acting. Second, the source of these sentiments is what the Vicar calls 'the cause of our nature.' Third, these sentiments are species-specific – 'suitable to our nature.' Fourthly, the sentimental promptings of conscience have an essentially self-directed moment: they concern the compatibility or incompatibility between me and the things in my environment and are grounded in a form of self-comprehension. Here, I submit, we find a first example of Rousseau's neo-Stoic naturalistic schema put to work in considering the role of self-consciousness in action. In this instance the adherence to Stoic doctrine is quite strict. We do well by following nature just as the animals do well by following nature. Why? Because nature (or 'the creator of our nature' – it comes to the same thing) has endowed us with a normative understanding of our own nature and what fits or befits it. If we allow ourselves to feel those promptings and to follow them we will be naturally guided to forms of action appropriate to us (and repelled from inappropriate actions), given the distinctive kind of being that we are. In the non-human animals we call this the prompting of instinct; in human psychology Rousseau calls it the 'instincts of the soul' or 'conscience.'

#### §4: Confessional Self-Consciousness

If only we were like the other animals, then conscience would be enough. We would find ourselves naturally attracted to what is good for us and naturally repelled from what is not. We could live well simply by following our instincts. We would not be alienated from our natural goodness and natural fit with our environment. But we are unique among the animals, whether for ill and for good. Our instincts have become suspect and unreliable, and our relation to our own nature is fundamentally complicated by our 'perfectibility' and by our freedom. A large and celebrated part of Rousseau's corpus is devoted to exploring the fact that we have ended up so wicked and ill and corrupt. I shall not here try to recount or assess the details of that Rousseauian story of the fall. Just take it as given that we have fallen. What are we to do then? A crucial part of Rousseau's answer, I want to propose, concerns a second and distinctively human form of self-consciousness, one that is deeply rooted in human nature, and which figures centrally

---

<sup>17</sup> For details of this analysis, see the draft of my paper on this topic online: <http://philosophy.fas.nyu.edu/docs/IO/2575/martin.pdf>. For a sampling of Stoic texts laying out the doctrine of *Oikeiosis*, see Sedley and Long, *The Hellenistic Philosophers*, I, 346-354.

in Rousseau's neo-Stoic conviction that our nature provides us with the means for the recovery of our moral health. Self-consciousness of this second variety I propose to call *Confessional Self-Consciousness*.

Here I come to another point in my investigation where I find myself swimming in alarmingly deep waters. The topic of confession obviously has a very rich and complex history in Western thought and culture: in the history of Christian theology and pastoral practice, of course, but also in the history of psychiatry, the history of literature, the history of the human sciences, and so on. Even within the comparatively narrow confines of Rousseau-studies there is a substantial literature on the theme of confession, which is a genre of writing which Rousseau is generally credited with having reinvented. In light of my rather spectacular ignorance about all this, I ought by all rights to keep quiet. But I propose instead to speak very tentatively to this theme by advancing five hypotheses for consideration and further investigation.

The first and least controversial point here concerns the importance of confessional autobiography in Rousseau's corpus and philosophical project. Over the course of his life Rousseau published three explicitly autobiographical works. The first and most famous of these is of course *The Confessions*, which were followed by *Reveries of a Solitary Walker*, and finally by *Rousseau: Judge of Jean-Jacques*. In all three cases it is fair to describe the mode of autobiography as confessional in at least the following sense: Rousseau does not confine himself to recounting the external events and episodes of his life, but seeks to expose to scrutiny the most intimate aspects of his private life and emotional state. He notoriously goes out of his way to explore his vices as well as his virtues, the events in his life which occasion shame as well as the accomplishments of which he is most proud. Looking beyond these explicitly autobiographical writings we find an autobiographical element in much of Rousseau's broader corpus as well, most notably in the Second Discourse, in *Émile*, and in the novels. In all these cases one of Rousseau's distinctive strengths and contributions is found in his subtle psychological observations, where he himself figures as the primary object of psychological study. To take just one example, the distinction between *amour-de-soi* and *amour-propre* quite clearly has an autobiographical element to it – born in no small part out of a characteristic self-scrutiny and self-interpretation of his own (often destructive) craving for recognition. If this is right then Rousseau not only contributed to the literary genre of confessional autobiography; his autobiographical self-examination itself occupies a central place in his philosophical project. Let me advance that as my first hypothesis.

In thinking about the significance of these autobiographical undertakings it is worth considering first their function. Unlike the sentimental self-consciousness that Rousseau finds at work in conscience, confessional self-consciousness is a rather elaborate action, deliberately undertaken. It is thus fair to ask about the purpose or purposes that such autobiographical undertakings are meant to serve. In short, what are these confessional writings *for*? If we can answer this question then we will also be in a position to say something about the conditions under which particular acts of confessional self-consciousness should be

counted as a success or a failure. This brings me to my second hypothesis, which at least in its general form I also take to be uncontroversial: confessional autobiography is purposive.

So what is its purpose exactly? I don't pretend to have a remotely adequate answer to this question. But we can get some leverage on it by considering what Rousseau himself says in the short prolog that he appended to the original manuscript of the *Confessions*.

Whoever you may be whom my destiny or my trust has made the arbiter of the fate of this notebook, I entreat you by my misfortunes, by your innermost emotions, and in the name of the whole human species not to destroy a unique and useful work which can serve as the first piece of comparison for the study of men, a study which certainly has not yet begun, and not to take away from the honor of my memory the only accurate monument to my character that has not been disfigured by my enemies.<sup>18</sup>

In Rousseau's memorable and characteristic plea we find a rather complex gesture that serves to exhibit a number of important aims of the confessional enterprise. Start from the final note in the plea: the complaint about having been 'disfigured' by the slander of his enemies. Throughout the autobiographical writings we find this theme of persecution: tale after tale of being subject to misfortune, misunderstanding and betrayal.<sup>19</sup> Rousseau's autobiography is undertaken in part to answer these disfigurements with a figure of his own. It should be clear, then, that one of the abiding aims of the autobiographical writings is a form of *self-justification*. Rousseau wants to tell his side of the story, *his* story, in order to justify himself both to his contemporaries and to posterity.

But in the plea of the prolog, the task of self-justification is subordinated to another. Rousseau is clear that he intends for *The Confessions* to make a contribution to science, in particular to the science of human beings. Rousseau resolves to expose every detail of his life -- public and private, virtuous or vicious, honourable or shameful -- in order to make a human life fully available for scientific study for the first time. This is not simply a matter of reading off the external events of a human life; a full confession will be one that exposes every corner and facet of the human soul. Rousseau held that this required an unprecedented self-examination and self-discovery, in order to provide what he calls "a self-portrait ... exactly according to nature and in all its truth". On this basis we can say that a second aim of the confessional undertaking is *self-knowledge* – in the first instance to provide Rousseau with the materials for knowing himself, and then in due course to provide the materials whereby human beings might finally come to a kind of collective self-knowledge of their own distinctive and complex nature.

---

<sup>18</sup> *Collected Works of Rousseau* 5:3. The prolog has a rather uncertain place with regard to the main text of *The Confessions*. Some editions do not publish it as part of the work, given that it seems in the main to make reference to the fate of the manuscript itself.

<sup>19</sup> It is worth remembering that the project of the *Confessions*, in particular, was undertaken in earnest in a period when Rousseau had been subject to intense attack, both for the content of his writings and for the conduct of his personal life. All three explicitly autobiographical works were undertaken in a period when there was an outstanding warrant for Rousseau's arrest in France, and in the case of *The Confessions*, in a period when Rousseau was living in exile. But it was not merely (or even primarily) to answer the legal charges that Rousseau takes up his autobiographical pen. Indeed there is comparatively little attention devoted to the legal allegations in any of these writings. What receives considerably more attention, particularly in *The Confessions*, is Rousseau's attempt to provide his side of the story regarding his high-profile falling out with his former friends and supporters.

There is an obvious tension between these two aims – between self-justification and a totally disclosive self-knowledge. It is far from clear that a complete fulfillment of the latter aim will in fact serve the former. (This became vividly clear at the first marathon public reading of Book I of *The Confessions*, after which one of Rousseau’s long-time supporters called for the police to prevent any further such readings.) At the same time, I want to argue, the aims of self-justification and self-knowledge are in fact two manifestations of a deeper common task: the task of *self-judgment*. As we have already seen, the word *judgment* figures in the title of the third and last of Rousseau’s autobiographical books: “Rousseau, *Judge* of Jean-Jacques.” The same word figures prominently in *The Confessions* as well – three times in the first dozen lines alone. The task of self-justification and the task of self-knowledge both presuppose this more fundamental undertaking of using confessional self-consciousness to reach some kind of judgment about myself and the life I am leading. This is accordingly my third hypothesis: for Rousseau, confessional self-consciousness is a form of self-judgment.

In the religious practice of confession there is a fundamental but complex relationship between the act of confession and the act of judgment. The nature of this relationship is far from uncontentious, but I hope it is safe to say at least this much: the confessant undertakes his confession in order to submit himself to judgment – not, I think, to the judgment of the confessor, but ultimately to the judgment of God. But the theology of this circumstance is quite complex. For of course I am said to be subject to the judgment of God whether or not I confess to my crimes. So the act of confessing cannot itself be what incurs that judgment. One might say that one confesses in order to ‘own up’ to one’s sins – that is, to acknowledge them as one’s own and to seek judgment. But even this cannot capture the theological significance of the action, insofar as the act of confession involves *renouncing* one’s sins, and since part of what one seeks from confession is a kind of *escape* from judgment, rather than judgment itself. I shall not here attempt to sort out these theological complexities, though I do think that they are significant for understanding Rousseau’s undertaking. Suffice for now to observe that Rousseau clearly has the religious paradigm in view, while also quite explicitly seeking to overturn it. That it lies in view is plain from Rousseau’s opening blast: “Let the trumpet of the last judgment sound when it will; I shall come with this book in my hand to present myself before the Sovereign Judge.” That he seeks to overturn it is evident from the act of refusal implicit in the title of the final work. For Rousseau neither God nor one’s friends nor one’s enemies nor posterity nor the reading public can serve as one’s confessor and judge; we are in the end judges of ourselves. With this we have come back to the problem of modern freedom, and to two final hypotheses which relate to this theme.

Fourth hypothesis: Confessional self-consciousness is therapeutic. I argued above that Rousseau’s naturalism is best understood as the modern application of a schema drawn from the Stoics. As we have seen, a central element of that schema lies in the conviction that we can find a natural cure for our moral ailments. Confessional self-consciousness, I propose, is intended to play a role in Rousseau’s neo-Stoic natural cure. So what exactly is its therapeutic role? We can characterize it in two different ways,

depending on how we think of the root causes of modern moral malaise. On one familiar Rousseauian diagnosis, our moral illness derives from the corruption of our instincts and values, particularly under modern conditions of civilization. Modern life has led to the emergence of a dangerous and fundamentally corrupt set of motives and sentiments, chief among them the inflamed *amour-propre* which in turn leads us into morally destructive forms of social dependency: we value appearances more than reality, we find ourselves only in the way we are represented by others. In this modern predicament, Rousseau holds, we lose the natural moral orientation provided by the self-sentiments of our conscience; worse: we learn systematically to repress them. We are left in a turmoil of confused and contradictory impulses and inclinations – some natural and healthy but weakened, others fundamentally corrupt and enormously potent.

If we accept this diagnosis, then the possibility of a return to moral health requires that we sort out this tangle of sentiments, learning to distinguish and cultivate those motives which are natural and noble while weeding out those which are artificial and base. For such an undertaking, confessional self-consciousness will have an important role to play. By undertaking a full confession I can hope to trace the origins of my sentiments and motives. With this genealogy and the neo-Stoic conviction about the goodness of nature I find myself with a therapeutic tool and a framework for moral progress. In the simplest version of Rousseauian ethics this is exactly the form that moral deliberation must take. As the Vicar has it, to live well I need only reawaken myself to the natural sentiments of conscience and resolve to follow their promptings.

But there is a second, and I think deeper Rousseauian diagnosis of modern moral ailments, and this diagnosis suggests a rather different account of the therapeutic function of confessional self-consciousness. To get at this second account we need only ask ourselves why we humans, among all the animals, are the ones so vulnerable to this alienation from our natural goodness. The answer, for Rousseau, goes back to human perfectibility and freedom. For the non-human animals, natural instincts and forms of life are fixed within very narrow boundaries. Neither koala nor salmon face any choice about how to act; both eat whatever presents itself as food, and their instincts are so-configured that (with a few fatal exceptions) only that which *is* food *seems* like food. But human beings are the ultimate omnivores. We can eat almost anything, and so unlike other animals have both the freedom and the task of choosing our dinner. What goes for diet goes for pretty much every thing else in human life. Form of life is simply dictated to the animal; for the human being it is posed as a question. For Rousseau it is this fundamental freedom in human existence which is the deep source of the danger of moral degradation discussed above. Because our form of life is not dictated to us we are capable of living in ways that in turn come to corrupt our natural moral orientation. Accordingly any adequate natural therapy must address itself to the fundamental paradox in the circumstance of the animal that is free by nature.

Here once again I want to propose that Rousseau's solution draws on Stoic leads, and that confessional self-consciousness plays a fundamental role in his proposed neo-Stoic solution. The key here

is to keep in mind Rousseau's neo-Stoic account of the workings of conscience. For Rousseau, as for the Stoics, conscience provides us with an innate normative orientation in our environment *by way of a distinctive form of self-representation*. This self-representation is not discursive or explicit, but practical and sentimental. Both animal instinct and innate human conscience are rooted in an implicit comprehension of *what kind of being one is*, and what suits or befits an entity of that sort. In Stoic terms, every organism is endowed with an innate sense for its own constitution or governing principle – the principle of mereological unity governing the proper integration of its spatial and temporal parts. But what kind of mereological totality is man – and in particular modern man? For Rousseau this unity, the unification of life-parts into life-whole, is not given but must be forged. And confessional self-consciousness, I propose, provides the medium for this distinctively modern human undertaking.

Final hypothesis: confessional self-consciousness is essential. By this I certainly do *not* mean to suggest that everyone must sit down and write a confessional autobiography in the manner of Rousseau. My claim, rather, is that an undertaking of that sort is in a certain sense part of our essence. At the outset of our discussion we found Rousseau complaining about the distortion of human nature wrought by the materialist naturalists of his day. Recall in particular the two human capacities that Rousseau found most deeply distorted by the materialist account: memory and judgment. Memory and judgment are the two psychological capacities most deeply implicated in confessional self-consciousness. In contrast to the natural self-sentiment at work in conscience, confessional self-consciousness requires that we first remember what we have done (and felt and thought ...) and then pass judgment upon ourselves. The judgment in question is not simply whether I have lived well or not, though such questions must surely arise as part of the confessional undertaking. But the deeper challenge is to judge *what kind of being I am*, and in particular what form of totality my life is to exhibit. If Rousseau is right, confessional self-judgment of this sort is part of the modern human essence. For we moderns are beings who are *essentially* – by our very essence – faced with the task of adopting a constitution. That is, we must find or discover or invent and articulate a form of totality that in turn gives meaning to our parts.

Confessional autobiography of the sort I have in mind can come in many different forms and media. As we have seen, Rousseau himself compares his task to that of the self-portraitist, who undertakes confessional self-representation in paint.<sup>20</sup> Confessional autobiography might take the form of therapy, or the kind of telling of one's life that so often plays a role in the forming of intimate attachments. In all these cases we undertake, as Rousseau undertook, to find sense in the shape of our lives by knitting the parts into the shape of a whole which in turn gives sense to the parts of which it is comprised.

---

<sup>20</sup> I discuss this form of self-portraiture in "Bubbles and Skulls: The Phenomenological Structure of Self-Consciousness in 17th Century Dutch Still Life Painting"; in Dreyfus and Wrathall (eds.) *A Companion to Phenomenology and Existentialism* (Oxford: Blackwell, 2006), 559-584. A text-only draft of this paper is available online: <http://privatewww.essex.ac.uk/~wmartin>

One distinctive possibility worth mentioning here – and with this I shall conclude – is the idea that this sort of confessional undertaking might be inscribed not in script or in speech or in print or in paint but in the medium of one’s very behavior. Suppose a young Marine Corp recruit spends his weekends riding one of those very high powered motorcycles along mountainous roads in the high desert, taking hairpin turns at sixty miles per hour. What shall we say of his behavior? We might say that here is another example of a young man in denial about his own mortality, failing to acknowledge or even actively denying the danger of his actions. We might bemoan the waste of the investment by the Department of Defense when so many such young men are injured or killed in motorcycle accidents before their expensive training is completed. But this would be to miss an important dimension of the action. What we see here, I propose, is an example of Rousseauian neo-Stoic Confessional Self-Consciousness at work – here in the medium of life itself. The young marine on his crotch rocket is representing himself for himself – representing the distinctive form of mereological unity that he has adopted. Here is a modern man who courts danger and risks his life in the course of his duties – not because he is simply following orders or because he has a contract with the Marine Corps to do so, or because this form of life has been passed down to him as an inescapable inheritance, but rather because he has chosen to do so. This activity within his life serves to reflect something about his life as a whole, and in so doing to give sense to the parts that compose it.

This sort of confessional undertaking, I have tried to argue, occupies a fundamental place in modern human life, and in the schema of neo-Stoic naturalism around which Rousseau’s ethics and moral psychology is organized. Like the animals, we live well by following our nature. But we human animals have a nature that presents us with a distinctive and paradoxical task. The constitution of human life – the distinctive form of totality that we ourselves are – is fundamentally indeterminate. It thus falls to us to discover and articulate a form of unity that can give sense and orientation in our endeavors. In the capacity for reflection and self-representation, and specifically in the capacity for what I have here called confessional self-consciousness, we find in our nature the means for confronting the paradox that nature has dealt us.