

Rousseau's Conscience in Modern Moral Philosophy

I. Introduction

1.1 While Rousseau's influence on modern moral philosophy has been widely acknowledged by recent commentators, that influence is often portrayed as indirect, through the role of the concept of autonomy in his political philosophy, or as merely inspirational, through his championing the moral self-sufficiency of the common man. For instance, John Rawls accords Rousseau a prominent role in shaping the tradition of political philosophy that locates the legitimate source of political authority in the individual's capacity to govern herself.¹ However, when Rawls' thought turns to foundational problems in moral philosophy Rousseau's identification of self-determination and morality does not seem to resonate.² Or consider Charles Taylor's and Jerome Schneewind's respective accounts of the development of modern moral philosophy, each of which restricts Rousseau's influence to the import of his thought that moral sensitivity does not derive from esoteric sources, whether they be religious or scientific, but rather is naturally available to the common man.³ According to these histories, it is Kant who reworks the naïve idea that morality is somehow contained in our nature into a wholly original understanding of the foundation of moral requirement. To both critics and enthusiasts, Kant is taken to be the originator of the idea that we understand how there might be actions that absolutely must or must not be done by grasping the idea of a free and self-determining agent. The present essay aims to correct a pervasive neglect of Rousseau's real innovation in modern moral theory: Rousseau had already, even before Kant, made the crucial "Kantian" move: he took conscience to be not merely a source of sympathetic motivation or a mere moral sensitivity, but rather the fundamental principle of

*Acknowledgements omitted.

¹ Rawls 1971, 264.

² Rawls 1971, 264. Here, in a footnote Rawls says that Kant gives a deeper reading of Rousseau's famous remark: "to be governed by appetite alone is slavery, while obedience to a law one prescribes to oneself is freedom." Rawls does not simply say that Kant develops the idea contained in this remark. Also see Rawls' treatment of Kant as the origin of the "idea that reason, both theoretical and practical, is self-originating and self-authenticating" (Rawls 1993, 100).

³ Schneewind 1998, 487-492; Taylor 1989, 355-363; Allen Wood suggests something similar in his introduction to *Immanuel Kant: Practical Philosophy* (Wood 1996, xvii).

the free human will.⁴ Appreciating the fundamental status of conscience for Rousseau will further enable us to see the availability of a metaethical viewpoint that may seem foreclosed to us just as it seemed to some of Rousseau's most illustrious contemporaries: a viewpoint according to which motivated action is an expression of our rational nature, where that nature is characterized not by its mysterious access to an equally mysterious realm of normative fact, but rather by its social constitution and autonomous operation.

1.2 We can introduce the problem of moral requirement by setting up a dilemma. It takes as its starting point Hume's thought that "the first virtuous motive, which bestows a merit on any action, can never be a regard to the virtue of the action" (T 478), and leaves one either with the metaphysical and epistemological excesses of rationalism, or with an empirical naturalism that cannot account for the necessity of the moral 'must'. This, of course, needs some explanation.⁵

Moved in part by a distaste for appeals upward and outward to a supernatural order to understand morality, the empirical naturalist instead turns his attention forward and inward towards human beings—to what they do and why they do it, to what they approve of and why they approve of it. It does seem to be a necessary condition of a "non-spooky" moral theory that it proceed in something like this manner; that is, from an understanding of what goes on in human beings, to an understanding of right and wrong. But we should note a potential equivocation in this expression "what goes on in human beings," what might be called the two senses of *action's ground*. On the first, when we ask about "what goes on in human beings" we are asking about something like the *structure* of human motivation; we are asking with Plato about the parts of the soul, or with Kant about the determining grounds of the faculty of desire, or again, and in a more contemporary spirit about kinds of explanatory reason. In contrast, on the second interpretation when we ask about "what goes on in human beings" we are asking about the *contents* of the various kinds of motivation. To see what I mean, suppose that we are all ultimately moved by sense appetite, it is still an open question

⁴ This essay contributes to Ernst Cassirer's understanding of Rousseau's place in the history of moral philosophy: See Cassirer 1951, Cassirer 1989, and Cassirer 1945.

⁵ See Korsgaard 1989 for such an account. My sense of how to frame the concerns driving modern moral philosophy has been influenced by this essay.

which appetites we have. Or suppose that it is settled with Hume that all rational motivation is ultimately based on passion, it is still an open question what passions move us: Are they all self-regarding? Is there a passion for cruelty? An inquiry into what goes on in human beings that aims at an understanding of morality must consider both and must not confuse them.

It is a starting point of Hume's "that all virtuous actions derive their merit only from virtuous motives, and are consider'd merely as signs of those motives" (T 478). His thought here isn't only that if a particular action is praiseworthy then it is done from a praiseworthy motive. It is also that motives are the primary object of moral evaluation; it is in terms of these that we are to understand which actions are morally correct. That is, Hume accepts a motivational analysis of morally correct action; such an analysis proceeds by asking which motives have moral worth and then asking which actions are to be done from such motives. The morality of motive and action are linked by definition: action *A* is morally correct if and only if *A* would be done from ideal moral motives. We have independent access only to the right side of the definition. The problem, then, is to determine what the ideal motives are.

Hume's famous argument that the "first virtuous motive, which bestows a merit on any action, can never be a regard to the virtue of the action," (T 478) can be read as ruling out the possibility of a particular motivational analysis—namely one which takes the motive of duty, or again the motive to do what is right, good, or virtuous as the way in. What Hume argues is that any such account will be viciously circular:

To suppose, that the mere regard to the virtue of the action, may be the first motive, which produc'd the action, and render'd it virtuous, is to reason in a circle. Before we can have such a regard, the action must be really virtuous; and this virtue must be deriv'd from some virtuous motive: And consequently the virtuous motive must be different from the regard to the virtue of the action. A virtuous motive is requisite to render an action virtuous. An action must be virtuous before we can have a regard to its virtue. Some virtuous motive, therefore, must be antecedent to that regard. (T 478)

The idea is that there is nothing for the motive of duty to latch onto, if duties are fixed by motives, or again that the principle "Do what is right because it is right" is, by itself, empty. It leaves one in the dark about which actions are right and so by itself can't direct one to perform an action or explain why one performed an action. The sort of doubt raised here is,

I think, an instance of what Christine Korsgaard has called content skepticism—skepticism about the bearing of rational considerations on deliberation and action.⁶ In this case it concerns the emptiness of the principle of duty or, as we might say in a more Rousseauian spirit, the apparent emptiness of conscience.

If “the ultimate object of our praise and approbation is the motive” (T 477) and the first moral motive can’t be a regard for the morality of the action, then, as Hume sees, “no action can be virtuous, or morally good, unless there be in human nature some motive to produce it, distinct from the motive of duty” (T 479). This sets the problem for the empirical naturalist—to determine what the distinct motives are. As I understand him, the empirical naturalist appeals to *thoroughly independent motives* here: most commonly self-love (modified by external constraint), or *sui generis* other-regarding psychological forces. The independence of such motives has two aspects: (i) the specification of their content does not employ any fundamental moral concepts, e.g. virtuous, right, or good; (ii), there is no need to look to a wider context than an individual bearer of the motive to understand its presence in her.

In Section II, I lay out Rousseau’s objections to empirical naturalism, focusing on this dual independence of its candidate motives. For now, however, I simply want to mention that the rationalist response to empirical naturalism was to say that it fails to account for the universality and necessity of moral requirement. According to the rationalist, this is because on such a view there is only a contingent relation between the candidate motives and what is intuitively regarded as right action, and also because there is no obligation to have the relevant motives and so no obligation to perform the relevant action.

In accord with the constraint of Hume’s “first virtuous motive” argument, the rationalist intuitionist (I’ll take Richard Price as a model) abandons the analysis of rightness of action in terms of motives: indeed, abandons trying to understand right and wrong by understanding what goes on in human beings. With Price, rightness becomes an intrinsic quality of actions that we have the power to perceive. To account for our intuition that moral judgment and motivation are intimately connected, Price simply asserts that awareness of the rightness of action is sufficient to motivate of duty; the first is stipulated and when rightness is an intrinsic property of actions it is possible simply to act because of it. But all this comes at a

⁶ "Skepticism about Practical Reason," in Korsgaard 1996, 311.

familiar cost. Price abandons an understanding of morality as an expression of our nature—the thought behind the Humean motivational analysis—and, I think, is at great risk of severing any plausible or intelligible connection between motivation and moral judgment. At the very least, the gulf that opens between judgment and action in such a view presents an obstacle to establishing their necessary connection. And where it is only an accident that the creatures which can intuit moral qualities are equipped to be moved in accord with that recognition it will be impossible to give a satisfying answer to the question “Why be moral?” There are, as well, familiar worries about the metaphysical credentials of moral qualities understood as primary qualities.

In short, either (a) we understand the morality of actions in terms of independently specified motives, or (b) we maintain that the first moral motive can be a regard for the morality of the action. If (a), then the motives in terms of which we account for the morality of actions must be other than a regard for the morality of the action, and then we come upon empirical naturalism and its problems. If (b), then we must abandon the thesis that the morality of actions is to be understood in terms of motives, and so come upon rational intuitionism and its problems.

1.3 I set up the question of the nature of moral requirement in this way because I take Rousseau and Kant to reject the empirical naturalist (Section II) and rational intuitionist (Section III) responses to the dilemma for similar reasons. Taking the first horn, the empirical naturalist cannot properly make sense of a moral requirement as applying universally and categorically – he cannot make sense of it as an object of knowledge. Taking the second horn, the rational intuitionist is committed to an unacceptable metaphysics and leaves the epistemological and motivational features of moral thought utterly mysterious – he cannot make sense of its efficacy. In rejecting both approaches, we can see Rousseau and Kant acknowledging that any adequate account of moral requirement will treat the motive of duty or conscience as that in terms of which we understand right action – the measure of acting well is internal to the will itself.

In other words, each refuses to buy into Hume’s dilemma. But by what right? What grounds do Rousseau and Kant have for maintaining that we are to find our way into

morality through an understanding of the motive of duty? To answer this is to look at the different ways in which Rousseau and Kant take themselves to be under a philosophical obligation to address the question of the apparent emptiness of the principle of duty (Kant) or conscience (Rousseau). Their approaches are markedly distinct. On the one hand, Kant rejects Hume's claim that the principle of duty is empty; he claims to be able to derive the supreme moral principle from common moral cognition and then show how it is that content, i.e., substantive duties, is contained in that that principle by reflecting on the universalization procedure. Rousseau, on the other hand, seems not to reject, but to ignore, Hume's challenge. In the absence of such a response to Hume's challenge one might reasonably suggest either that Rousseau's account is a failure or take that failure to be a demand for a different interpretation of his moral theory. But, I'll argue (Section V-VI) that the absence should not tell against an interpretation of Rousseau's moral theory on which conscience is the fundamental principle of the free human will. Instead, his silence on this count is driven by moral considerations: he doesn't attempt an analysis of conscience because conscience forbids it. This point reflects a deep feature of Rousseau's philosophizing that puts him somewhat at odds with much contemporary practice: the conviction that philosophy at its heart is as much an ethical as an epistemic enterprise, and perhaps an ethically dubious one at that.

II. Against Empirical Naturalism: "Flee those who sow dispiriting doctrines in men's hearts under the pretext of explaining nature." (E 312)

1. Rousseau's objections to empirical naturalism take a variety of forms, some commonplaces of his period and others more distinctive and innovative. I begin by considering what might be regarded as a moral objection to the sentimentalist construal of acting from duty. Then I turn to the problems that stem from what I called above the thorough independence of empirical naturalism's candidate motives. Rousseau holds that these empirically given motives – motives other than duty, conscience, etc. – bear only a contingent relation to right action. As I will characterize this familiar objection, it is directed towards the *content* of the favorite motives of the empirical naturalist. Bearing in mind our two senses of *action's*

ground, I then develop a different and perhaps less appreciated line of Rousseau's thought. Here, the empirical naturalist is charged with being unable to explain the essentially relational character of morality – or at least that part of morality which has to do with what Kant calls the “practical relation of one person to another, insofar as their actions, as deeds, can have (direct or indirect) influence on each other” (MM 230)⁷, i.e. private right. The basis for this charge, in my reading, is already contained in Rousseau's well-known criticism of Hobbes in the *Discourse on Inequality*. As I will characterize it, this more metaphysical objection is directed towards the empirical naturalist's understanding of the *form* of moral motivation.

2. A familiar objection to any attempt to understand morality in terms of motives other than the motive of duty argues that there is only an accidental connection between the candidate motive and what common moral cognition takes to be right action. And Rousseau avails himself of this point. I'll consider the two broad classes of motive available to the empirical naturalist: self-love or *sui generis* other regarding sorts, in turn.

Against the theory that self-love is the ultimate ground of moral motivation and evaluation, Rousseau advances some standard sentimentalist arguments. According to Rousseau, the self-love theory can't explain why moral approval or disapproval attaches to actions or characters which have no relation to the appraiser's private interest: “it is surely of very little importance to us that a man was wicked or just two thousand years ago; nevertheless, we take an interest in ancient history just as if it all had taken place in our day” (E 288). Even more problematic, the *experimentum cruces* as Hume calls it in his discussion of the same problem (ECPM 42), is whether the self-love theory can plausibly explain how one might be motivated to do what couldn't possibly be seen to be in one's interest. Rousseau asks with Hume, “What is going to one's death for one's interest?” (E 287). The self-love theory, Rousseau decides, is “too abominable a philosophy – one which is embarrassed by virtuous actions, which could get around the difficulty only by fabricating base intentions and motives without virtue” (E 289).

⁷ Translations of Kant's *Metaphysics of Morals* and *Groundwork for the Metaphysics of Morals* are Mary Gregor's, while page numbers refer to volume 6 of the *Akademie-Ausgabe*.

The empirical naturalist regards the problem of the origin and foundation of morality as primarily a psychological problem. So, not surprisingly, the sentimentalist response to the explanatory deficiencies of the self-love theory itself travels at the level of empirical psychology. It recommends an alteration in our understanding of the original materials of human concern and motivation. Usually this involves adding a sympathetic disposition – a “fellow-feeling or humanity,” as Hume calls it, which is in all and directed towards all – as well as other more specific items, e.g., a father’s natural affection for his child (T 478). It seems likely that such a maneuver is sufficient for accommodating the above cases in which evaluation or action is seemingly distanced from a subject’s own well-being. Indeed, Rousseau is friendly to the idea that man has an innate repugnance to see his kind suffer (DI 160) – going so far as to claim that “men would never have been anything but monsters if Nature had not given them pity in support of reason” (DI 161). He nevertheless does not regard the sentimentalist response as adequate. For Rousseau, sympathy – conceived as a basic and natural ability to feel and respond to the suffering of others on the basis of sense perception (DI 162) – is not a moral capacity, and certainly not the central moral capacity. But why?

One reason for rejecting a morality of sympathy is that sympathy, like self-love, is only accidentally hooked up with right action. As Kant would later argue in the Preface to the *Groundwork*:

in the case of what is morally good it is not enough that it conform with the moral law but it must also be done for the sake of the law; without this, that conformity is only very contingent and precarious, since a ground that is not moral will indeed now and then produce actions in conformity with the law, but it will also often produce actions contrary to the law. (G 390)

Rousseau invokes this very contingency in opposition to the self-love theory when he tells us:

he who keeps his promise only for profit is hardly more bound than if he had promised nothing, or, at most, he is in the position to violate it like the tennis players who put off using a bisque only in order to wait for the moment to use it most advantageously. (E 101)

It is only an accident if a promiser who is guided by self-love has sufficient reason to do what he promised to do. And clearly there is a similar problem for a merely sympathetic agent since there are cases in which sympathy would have one break a promise when common

moral cognition would have one keep it, or worse cases in which sympathy would have one kill one to save five.

To be sure, the ethical naturalist may be cognizant of the fact that action from self-love or, say, sympathy can be contrary to what we pre-theoretically regard as right. Such theorists have very often appealed to a convention or a practice to effect the appropriate, strict and general kind of connection between their favorite motive and right action. Pursuing the dialectic at this level, though, misses the deeper critique that Rousseau is aiming. Appeal to a convention won't account for what is a plain fact for Rousseau—namely that we have natural duties and rights—duties and rights the having of which doesn't depend on being hooked-up to other agents through any particular convention or practice. For Rousseau, moral relations enter into all human relations; they do not presuppose that the agents enter into any particular conventional relations. To have knowledge of morality is “to have a sense of the true relations of man, with respect to the species” (E 219); it is to conceive of oneself as “an integral part of [one's] species” (E 220). The attempt to shore up a particular independent motive through appeal to convention thereby demonstrates, in Rousseau's eyes, a fatal commitment to independence in our second sense, a failure to grasp the foundational place of relationality in our ethical life.

3. According to Rousseau, “No good action is morally good except when it is done because it is good” (E 104), and the empirical naturalist, as I've construed him, must disagree with this. It is a familiar objection to the self-love theory that it provides a morally repugnant account of moral motivation. We can see a further elaboration of Rousseau's attack on the independence of motives in his objection to the Humean account of the motive of duty.

Although Rousseau does not direct his substantive critique of modern moral life and society at sentimentalism, it is implicated in that critique just where it treats the motive of duty as secondary. In the hands of the sentimentalist, acting on the motive of duty is best understood as acting from the desire for approval from oneself or others. If it is thought that actions derive their moral status from our approval or disapproval of them, it seems that our motivation to act for the sake of that status would also be derived from a desire for approval. Consider Hume's view of the matter:

But may not the sense of morality or duty produce an action, without any other motive? I answer, It may: But this is no objection to the present doctrine. When any virtuous motive or principle is common in human nature, a person, who feels his heart devoid of that principle, may hate himself upon that account, and may perform the action without the motive, from a certain sense of duty, in order to acquire by practice, that virtuous principle, or at least, to disguise to himself, as much as possible, his want of it. (T 479)

For Rousseau, this account of doing something because it is right is no mere philosophical confusion – no innocent conception of moral motivation. Rather, it inextricably binds our moral lives to duplicity. At the beginning of the *Discourse on Arts and Sciences*, he writes,

Today, when subtler inquiries and a more refined taste have reduced the Art of pleasing to principles, a vile and deceiving uniformity reigns in our morals, and all minds seem to have been cast in the same mold: constantly politeness demands, propriety commands; constantly one follows custom, never one's own genius. One no longer dares to appear what one is; and under this personal constraint, the men who make up the herd that is called society will, when placed in similar circumstances, all act in similar ways unless more powerful motives incline them differently. (DAS 6)

So the requirement that duty can only be a secondary motive entails a fundamental duplicity: so much the worse for our moral self-regard, a tough-minded thinker might respond. But Rousseau's criticism need not end here; without another way of understanding acting from a sense of duty, it will be impossible to treat morality and freedom as compatible. After all, the desire for mere approval is the source of our duplicity and is also, according to Rousseau, the source of slavish dependence,

For one's own advantage one had to seem other than one in fact was. To be and to appear became two entirely different things, and from this distinction arose ostentatious display, deceitful cunning, and all the vices that follow in their train. Looked at in another way, man, who had previously been free and independent, is now so to speak subjugated by a multitude of new needs to the whole of Nature, and especially to those of his kind, whose slave he in a sense becomes even by becoming their master. (DI 180)

But without being able to treat morality and freedom as compatible, Rousseau could argue that Hume also cannot treat morality and happiness as compatible – the aim of section 9, part 2 of his *Enquiry*.⁸

⁸ What are Rousseau's grounds for dismissing the moral motive as a motive for self-approval? There is no problem with *self*-approval as such. After all, Rousseau understands the rewards and punishments of the afterlife

4. For Rousseau, then, attempting to understand the motive of duty as derivative of the independent motive of other-approval leads Hume theoretically astray. But the situation turns out to be even worse for the empirical naturalist. The thoroughgoing insistence on independence turns out to leave him without a grasp on the conditions under which someone might be moved by such a passion in the first place.

This objection is developed via Rousseau's well-known criticism of Hobbes. But this criticism can prove significantly more slippery than many commentators appreciate. What exactly exercises Rousseau about the Hobbesian state of nature? Charity alone requires that we not regard Rousseau as merely complaining that Hobbes has given us an unacceptably nasty and brutish picture of man – such moralizing would be too obviously point-missing. We must also not accept a frequent classroom depiction of Rousseau's charge, according to which Hobbes mischaracterizes the state of nature so that "man is naturally intrepid and seeks only to attack and to fight" (DI 143), thereby permitting a construction of the state of nature as a state of war, and thereby setting the stage for a political philosophy which sees the individual's subjection to another as a necessary condition of her felicity. Putting the dispute in terms of the correct depiction of the state of nature prevents us from getting to the heart of the matter because "state of nature," rather than enjoying an independent existence that either theorist might succeed or fail in depicting correctly, functions as a technical term in their respective systems. For Hobbes it picks out a condition in which "men live without a common power to keep them all in awe" (L, I. 13.8). It is life without the state. The sense Rousseau attaches to the term is considerably more elusive; roughly, the state of nature is the condition in which none of man's developed capacities are traceable to "the faculty of perfecting oneself; a faculty which, with the aid of circumstances, successively develops all the others" (DI 149).

in terms of "the pure delight born of satisfaction with oneself and the bitter regret at having debased oneself" (E 284). However, in Hume's case, the relevant self-approval requires self-deception; the agent acts from the motive of duty "at least, to disguise to himself, as much as possible, his want of it [the natural motive]" (T 479). Moreover, the motive which is thought to be lacking is thought to be so merely against the background of its presence in the motivational set had in common by most others.

Rousseau's explicit complaint is that Hobbes "spoke of Savage man and depicted Civil man" (DI 139) when "he improperly included in Savage man's care for his preservation the need to satisfy a multitude of passions that are the product of Society" (DI 160). Famously, Rousseau's preoccupation is with glory. But why is it so clear that glory is a social motive? How are we to distinguish savage from social motives in the first place? And what prevents Hobbes from accommodating Rousseau's point without significant alteration to the rest of his theory? After all, aren't competition and diffidence, in the absence of glory, sufficient to make the state of nature a state of war? If so, why accord Rousseau's charge the real significance he seems to think it deserves?

5. In "The Natural Goodness of Humanity", Joshua Cohen suggests that we see Rousseau concluding that the relevant Hobbesian motives are social on the basis of the following three claims: (i) "cognitively complex passions require 'enlightenment'," (ii) "enlightenment is a result of social interdependence," and (iii) "the motivations relevant to Hobbes' account of human conflict are *cognitively complex*."⁹ According to Cohen, Rousseau makes a distinction between natural motives and concept-dependent motives:

By contrast with hunger or primitive sexual appetites or the desire for sleep, all other desires are concept- and belief-dependent. They depend in particular on opinions and judgments, and require that the subject be able to represent the particular object of desire as an individual with certain general properties.¹⁰

So, the appetites, more or less, are the natural motives; the naturalness of these consists in the fact that it is not a necessary condition of being moved by or having such a motive that an agent be able to employ concepts in the service of action. A creature might be moved by hunger even when it can't, say, represent its hunger as "a reason for acting." As these motives frequently find expression in *us* – for example, in response to the question "Why are you *A*-ing?" one might say "because I'm hungry" or "because I like to *A*" – they are evolved from forms of expression where we have instead to do with something on the order of groans and moans. Concept-dependent motives, on the other hand, are those for which the ability so to employ concepts is a necessary condition. Here the primary form of expression is linguistic –

⁹ Cohen 1997, 115.

¹⁰ Cohen 1997, 113.

for example, in response to the question “Why are you *A*-ing?” one might say “in order to *B*.” In order to explain oneself in this way and in order to have the concept-dependent motive of doing *B*, one must have the concept of *B*-ing, or grasp in thought what it is to do *B* – obviously I can’t intentionally do *B* or intend to do *B*, if I don’t know what it is to do *B*.

Cohen’s argument is that the conceptual capacities presupposed in motivation by a concept-dependent motive are essentially tied to facility with a language, and the existence of and facility with a language depends on the fact of coordinated interaction with others, i.e., social interdependence. If this is correct, reasons Cohen, then all concept-dependent motivation is “social.”

There is, undoubtedly, an interesting and important distinction between natural and concept-dependent motivation. But if Cohen’s reading of Rousseau’s distinction between savage and social motives were the right one, Hobbes would have no problem accommodating the objection. Although not restricted by positive laws in their interactions with others, inhabitants of Hobbes’ state of nature, whose practical rationality is perhaps captured by contemporary game theory, can nevertheless interact. But what could restrict Hobbes from holding that they can interact sufficiently often and with sufficient purpose so that various forms of coordinated behavior develop, e.g., linguistic practices. It does not follow from the fact that Hobbesian beings are such as to find themselves in prisoner’s dilemmas in the state of nature that they can’t also find themselves in non-competitive game theoretic situations. Indeed, Hobbes seems to suggest that there are such when he considers the place of confederacies there (L I. 13.1 and 3). Insofar as it is not unreasonable to understand a language as a practice whose origins can be modeled as a coordination problem, it is not unreasonable to hold that Hobbesian beings can be language users. Furthermore, I see no reason for thinking that Rousseau rejects the possibility of Hobbesian language users, and even see some evidence that he already accepts it (DI, 172-3). So, Hobbes could claim that while inhabitants of the state of nature don’t, for example, naturally have a concept-dependent motive to eat apples because they don’t naturally or innately have, the concept ‘apple’, they can develop this motive even while remaining guided by self-love and outside the structure of civil society. If so, agents in Hobbes’ state of nature might be moved by

social motives as these are understood by Cohen. This seems reason enough to seek a different account of Rousseau's distinction.

Granting inhabitants of Hobbes' state of nature a language and concept-dependent motives does not commit Rousseau to, as it were, granting unlimited scope to that language and those motives. For example, he need not admit that it has room for the concept of glory or that the speakers of it be capable of being moved in any way that *we* ought to describe as expressive of glory seeking.

Consider the following cases of apple-involving action:

1. Young Kaspar Hauser is eating what is in fact an apple because he is hungry.
2. Helen is eating an apple in order to gain three pounds.
3. Helen is grabbing the apple because, as she says, she herself is "the fairest of them all."

Now, since Cohen identifies the set of savage motives with the set of natural (or concept-independent) motives, he would have us treat 1 as the only case of savage motivation and would have us treat both 2 and 3 as cases of social motivation. But this is too coarse. The distinction important to Rousseau is not that between 1 and (2 and 3) – a distinction between motives which don't and do have living *among others* in their causal past – but that between (1 and 2) and 3 – a distinction between motives which don't and do have living *with each other* in their nature (E 235 fn.).

Why does Cohen take the set of social motives to be as wide as the set of concept-dependent motives? One assumption that may lead him to lump cases 2 and 3 together is the thought that, if one has the capacity to be moved by some concept-dependent motive, then one has the capacity to be moved by any concept-dependent motive. Were this true, then we could infer that some creature has the capacity to be moved by considerations of glory and justice on the basis of its being moved by an intention to eat an apple. It is just a question of adding a few more concepts. What might lead someone to accept this thought? Let me venture a further hypothesis: if one sees every concept-dependent motive as completely described by a suitable substitution for φ in "I want to φ " (or p in "I want that p ") then one will think that if one can have some substitution instance as a motive then one can have any substitution instance. All that is needed to meet the cognitive complexity requirement for having a motive is acquaintance with the relevant specification of φ . This, it seems, might be

done in the most ordinary of ways. According to this view, we fully understand a motive when we understand what an agent aims at who has it at all; all differences between concept-dependent motives are differences in content or matter.

Such a view may seem simply to encapsulate common sense about concept attainment. But let's look more carefully at the case of glory. If having the motive of glory is nothing but a matter of grasping a particular content, what might that content be? Here is a plausible candidate: to have the glory motive is to want to bring it about that another grant that one is better in comparison. But if wanting this is what it is to have the glory motive, how are we to distinguish case 3 above from the following:

3a. Private Haroldson races past his platoon in order to take out the well fortified bunker in order to make the other privates think comparatively little of themselves so that they will collectively give him the only weekend pass.

That is, how do we distinguish someone who *really* acts from the motive of glory and someone who only aims at bringing about the relevant state of affairs as a means to some further end? After a few iterations of such maneuvers, one will probably be driven to say that glory is wanting that another think little of himself in comparison *in the right way*. But this is just a concession that there is no non-circular specification of the glory motive in terms of its content. As a result, there can be no guarantee that one, simply by virtue of general conceptual competence, enjoys the capacity for being moved by or attributing the motive. It may well be that only those embedded in a way of life within which the concept makes sense are susceptible to the glory motive.

Someone might object that a suitable specification can evade examples such as that of Private Haroldson: glory is wanting that another think comparatively little of himself as a final end or last end. But consider a parallel attempt to reduce a natural motive like hunger to wanting to eat as a final end. On this suggestion we are to distinguish being moved by hunger and being moved by an instrumental desire to eat (say to avoid being hungry later) according to whether the motive is final. However, someone might want to eat for no further reason, and if so then wanting to eat is last in a series – it explains why I intentionally extend my arm and lift the cashew off the counter – without being the same as hunger. So, we come upon the problem of distinguishing talk of 'wanting to eat' when it expresses hunger and

when it expresses something I want to do for no reason. Wanting to bring it about that another think little of himself in comparison when it is last or final need not be the glory motive: it might also be understood as *merely* wanting to do that, i.e., wanting to do it but for no reason.

On Cohen's view the only interesting division among *forms* or *kinds* of motivation and agency is that between brutes and self-conscious concept exercising creatures like us. My suggestion is that just as Rousseau wants us to see a difference between the form or kind of willing, acting, and agency in play when Kaspar Hauser is eating because he is hungry and when Helen is eating in order to gain three pounds, so too Rousseau wants us to see a difference in the nature of the willing, acting, and agency in play when Helen is eating in order to gain three pounds and when she is grabbing for glory.

6. The way into the distinction between savage and social motives and the different forms of agency that underlie these is through Rousseau's distinction between *amour de soi* and *amour propre*. In Rousseau's system the power of desire or the generic capacity to act is oriented by either *amour de soi* or by *amour propre*; every motive, correspondingly, is grounded in either *amour de soi* or *amour propre*. Savage motives are based on *amour de soi*: "a natural sentiment which inclines every animal to attend to its self-preservation" (DI 226) and well being, i.e., pleasure (DI 132). Social motives are based on *amour propre*, a sentiment that inclines every animal which has it to rank itself in its own species (E 235, 278). The determinate expression of *amour de soi* will vary from species to species and also within some one species according to differences in surroundings and natural endowments (DI 171). What is less clear, or at least has been less clear to some interpreters of Rousseau, is that the expression of *amour propre* in human life can take a number of determinate shapes.

The first glance he casts on his fellows leads him to compare himself with them. And the first sentiment aroused in him by this comparison is the desire to be in the first position. This is the point where [*amour de soi*] turns into *amour propre* and where begin to arise all the passions which depend on this one. But to decide whether among these passions the dominant ones will be passions of beneficence and commiseration or of envy and covetousness, we must know what position he will feel he has among men, and what kinds of

obstacles he may believe he has to overcome to reach the position he wants to occupy. (E 235)

But what underlies these different shapes of *amour de soi* and *amour propre*? What are the common cores?

As they figure in developed humans, *amour de soi* and *amour propre* are kinds of rational self-concern. Rational self-concern is distinctive in that what one is concerned about is partly determined by one's conception of the kind of thing for which one has concern. So, by looking to Rousseau's characterization of the self-conceptions underlying *amour de soi* and *amour propre*, we are at the same time getting a glimpse of Rousseau's characterization of the different forms of agency, or orientations of will, involved in or underlying the savage and social motives. On the one hand, a human being guided by the absolute sentiment of *amour de soi* "views himself as the only Spectator to observe him, as the only being in the universe to take any interest in him, as the only judge of his own merit" (DI 226). On the other hand, a human being guided by the relative sentiment of *amour propre* views himself essentially in comparison with others; "as soon as *amour-propre* has developed, the relative *I* is constantly in play, and the young man never observes others without returning to himself and comparing himself with them" (E 243).

In denying that we can understand glory as merely a special content, of a sort that is in principle available to any being capable of concept use *überhaupt*, I am saying that Rousseau's idea isn't simply that when *amour propre* comes on the scene the solitary-I of *amour de soi* comes to have an additional and strong interest, want preference, desire, or member of its subjective motivational set, in the way that it would were the solitary-I to develop a habitual longing for, say, sun bathing or eating gummy candy. Rousseau's idea isn't that whereas the solitary-I doesn't care what others think about her, the relative-I does just happen to care about this. His idea is that with the development of *amour propre* agency itself takes on a radically different form. What was once a solitary-I is now a relative-I. A full account of the possibility and significance of this transformation lies beyond the scope of this paper. For our purposes, two lessons are paramount: first, the necessity for proposing such an alternate conception of agency, one according to which the availability of certain conceptual forms goes hand-in-hand with the possibility of certain forms of action and self-

understanding, arises in part from the great difficulty of explaining action in pursuit of glory in terms of a possession of a desire with a certain special content; second, this unsatisfactory explanatory strategy is precisely what the empirical naturalist is committed to by virtue of his twin commitments to independence: the conceptual independence of motives according to which ethical motives must be replaceable by non-moral specifications, and the independence of subjects according to which the motivation of the moral subject must be in principle abstractable from larger contexts of sociality.

III. Resisting Rational Intuitionism: "The entire right of nature is only a chimera if it is not founded on a natural need in the human heart." (E 235)

1. As I said above, the rational intuitionist gives up the motivational analysis of right action to protect the reality of obligations.¹¹ Laying out the threat of empirical naturalism Price says,

if no actions are, *in themselves*, either right or wrong, or any thing of a moral and obligatory nature, which can be an object to the understanding it follows, that, in themselves they *are* all indifferent ... But are we not conscious, that we perceive the contrary? (RP 147)

In order to accommodate our awareness that we give voice to necessary truths when we attribute rightness or wrongness to actions, Price asserts that right and wrong are "real *characters* of actions. They must immutably and necessarily belong to those actions of which they are *truly* affirmed" (RP 148). Price then equips us with a rational "power *immediately* perceiving right and wrong" (RP 142) to explain how we know which actions are right and which wrong. Finally, to account for our intuition that judging right or wrong and being motivated are intimately connected, Price asserts that "when we are conscious that an action is *fit* to be done, or that it *ought* to be done, it is not conceivable that we can remain *uninfluenced*, or want a *motive* to action" (RP 194). Price is thereby able to preserve both the

¹¹ I'll be taking Richard Price's *Review of the Principal Questions in Morals* as the model of such a view.

reality of obligations and the purity of the motive of duty: the first is stipulated and when rightness is an intrinsic property of actions it is possible to act straightforwardly because of it.

There are two pressures to interpret Rousseau as a kind of rational intuitionist. The first is philosophical and stems from recognition of his objectivism, while the second is textual, arising from a number of passages which appear to commit him to such a view. I'll address the former in this section (III) and the latter in the following (IV), and in each case argue that we have not been given reason to accept such an interpretation.

2. Rousseau is perfectly clear about the absence of moral laws in the state of nature (DI 131-3). His central complaint against those who include such laws stems from the thought that one ought not "to make a Philosopher of man before making a man of him" (DI 133, E 290). What Hobbes saw when he "very clearly saw the defect of all modern definitions of Natural right" (DI 159) was that "the desires and other passions of man are in themselves no sin. No more are the actions that proceed from those passions, till they know a law that forbids them" (L I. 13.10). But if there is no law with moral content that can be grasped by savage man, then what he does is not a possible subject of moral evaluation. For a law "to be natural, it must speak immediately with the voice of Nature" (DI 132), but conscience, the moral voice, does not speak to savage man who is concerned only with his own well-being.¹² But this does not decisively refute the intuitionist reading of Rousseau: it might be that while Rousseau thinks that man is not obligated in the state of nature because "man can be punished only for the mistakes of his will, and...an invincible ignorance could not be imputed to crime" (E 258), he nevertheless thinks that moral laws are there, in some sense. Savage man would stand to those laws as, for example, my fish stand to the laws of the State of Massachusetts.

¹² In my view, showing that moral laws are not part of Rousseau's state of nature is sufficient for demonstrating that Rousseau doesn't think that intuitionist "metaphysical principles" are *in* nature. But a full demonstration of this would require an appropriate account of why Rousseau uses a device such as the state of nature to consider the character of requirements on the will, moral or otherwise. I take the alternative tack through Rousseau's objectivism, in part because I believe it casts more light on both the attractiveness and the shortcomings of intuitionism, ultimately enabling a fuller appreciation of what makes Rousseau's notion of conscience distinctive.

The intuitionist interpretation can seem attractive through attention to Rousseau's objectivism; the guiding thought would be that if the moral law is to be an objective measure of the operations of particular creatures, then its validity cannot depend on such creatures. It must stand over against them, and so Rousseau must locate the standards of action elsewhere, namely, in the nature of things. Now, it seems correct that the fundamental measure of right action does not depend on the existence and operations of any particular people; if tomorrow each of us, or if in 150 years each of them, signed on to the Calliclean program of how to live, it would still be true that an unjust man is a bad man, an act of injustice a bad act. But this thought of the independence of moral law from the particular creatures subject to it, or again from human *beings*, must be distinguished from the thought that the moral law is constitutively independent of human *being* simply, or again what it is to be human. The one who assimilates Rousseau to Price simply on the basis of a commitment to the objectivity of moral requirement runs together these conceptions of the independence of moral standards and what it is a standard of, failing to see that the unproblematic independence of norm and individuals of a kind need not entail the far more fraught metaphysical independence of norm and the kind itself.

In order for Rousseau's objectivism to be sufficient for interpreting him as an intuitionist, it must be that there is no other way to account for this than by sticking moral requirements in nature. Now, Charles Taylor does think that there are other ways:

It is quite possible to conceive that the best theory of the good, that which gives the best account of the worth of things and lives as they are open to us to discern, may be a thoroughly realist one – indeed, that is the view I want to defend, without wanting to make a claim about how things stand for the universe 'in itself' or for a universe in which there were no human beings.¹³

But Taylor doesn't think Rousseau avails himself of such alternative possibilities. On Taylor's interpretation, Rousseau "ran his inner voice in tandem with the traditional way of understanding and recognizing universal good."¹⁴ The Deist tradition to which Taylor refers locates the good within the order of nature as instituted by God: we know what is required by seeing which actions fit into the order as it is discerned by reason, and we want to do

¹³ Taylor 1989, 257.

¹⁴ Taylor 1989, 361.

what is so recognized because of a natural love of what reason recognizes to be good, i.e., conscience. The picture is one in which the virtuous person has a faculty of desire that accords with knowledge of the universal good acquired through an entirely independent faculty.

Taylor contrasts Rousseau's position with what he takes to be its natural development in Kant : "the inner voice of my true sentiments *define* what is the good: since the élan of nature in me *is* the good, it is this which has to be consulted to discover it."¹⁵ But this is precisely the position Rousseau occupies. It is one of his distinctive and fundamental ideas that morality is not to be understood by looking to nature *tout court*, but rather, to *our nature*:

There is in the depths of souls...an innate principle of justice and virtue according to which, in spite of our own maxims, we judge our actions and those of others as good or bad. It is to this principle that I give the name *conscience*. (E 289)

It is because he views moral requirements as rooted in our nature as expressed in conscience that Rousseau does not think that the fact that "the idea of right... and still more that of natural right, are manifestly ideas relative to the Nature of man" (DI 131) is a threat to the objectivity of moral requirement.

3. Perhaps we can explain Taylor's urge to treat Rousseau on the model of intuitionism by appeal to the fact that the intuitionist and Rousseau each endorse a picture of the moral agent as cognitively self-sufficient and naturally good. Like the intuitionist, Rousseau holds that we need not rely on another to apprehend moral truths, and that recognition of the rightness of an action can, in some sense, prompt performance. The first thought is exhibited in Rousseau's frequent charge that moral knowledge need not, indeed cannot, be esoteric: "We can be men without being scholars" (E 290); "Either he will learn these duties by himself, or he is excused from knowing them" (E 303). The second is implied by his claim treated above that "No good action is morally good except when it is done because it is good" (E 104).

¹⁵ Taylor 1989, 362.

But is this sort of independence and natural goodness sufficient to account for the connection Rousseau seems to draw between freedom and morality? Is this conception of moral self-sufficiency enough to explain Rousseau's thought that we most fully express our nature as "active and free" when acting from conscience (E 380-1), or that "the sentiment of my freedom is effaced in me only when I become depraved and finally prevent the voice of the soul from being raised against the law of the body" (E 380)? Taylor seems to think so. According to him, it is because conscience is *within* us and because it is sufficient to motivate that Rousseau identifies morality and freedom. He says, "The distinction of vice and virtue, of good and depraved will, has been aligned with the distinction between dependence on self and dependence on others. Goodness is identified with freedom, with finding the motives for one's actions within oneself."¹⁶

In order for the thought that one is free and virtuous in "finding the motives for one's actions within oneself," to be illuminating, we need to understand the significance of a motive's coming from *within* and of an *inner* source of motivation. When I want, for example, to coach the Boston Red Sox or to build a stadium for them, there is something "external to me determining me" (E 380). In order to have such ends I must have "acquired knowledge" of them (E 286) and done so through the senses, i.e., that which puts us in touch with the external world. So it seems that being moved by such ends is quite different from being moved by that "faculty called instinct, which appears without any acquired knowledge to guide animals toward some end" (E 286). Of course, only animals, i.e., beings which sense, have instincts of this sort; nevertheless, these instincts are somehow already in them prior to the actual operation of the senses. Drawing on a traditional metaphor, Rousseau directs our attention to another sort of instinct, "divine instinct" or conscience, which he claims is "to the soul what [animal] instinct is to the body" (E 286). Rousseau seems to regard both natural needs, i.e., the appetites (DI 142), and conscience as inner sources of motivation. Yet, for Rousseau it isn't correct to say that one's freedom or true nature is expressed in the appetitive pursuit of food, water, sex or sleep, while it is correct to say this of someone who acts from one's conscience (E 380).

¹⁶ Taylor 1989, 361.

In other words, “coming from within” and “inner” cannot do the work Taylor wants them to do in characterizing Rousseau’s views. What he needs to account for the connection between morality and freedom is a way of distinguishing natural needs from conscience as inner sources of motivation. Taylor might say that Rousseau identifies the *real self* with the soul, pointing to such passages as the following: “man lives only halfway during his life, and the life of the soul begins only with the death of the body” (E 283); and “I aspire to the moment when, after being delivered from the shackles of the body, I shall be *me* without contradiction or division” (E 293). And Taylor might then say that only motives that arise from the soul *really* originate within an agent and that conscience clearly is such a source; it is after all the voice of the soul. But in order to see whether Taylor is entitled to this move, we must consider what drives Rousseau to talk of a soul in this context. Only then will we be in a position to ask whether, why, and in what way Rousseau takes conscience to be a part of it.

4. In dividing body and soul, Rousseau is not expressing some primitive intuition that certain things are immaterial and others material. Rather, he is led to this divide in the first moments of the “Profession of Faith” with the question “Who am I?”¹⁷ In asking who he is, Rousseau does not aim at an account of his idiosyncrasies or a catalogue of the details of his life. Instead, this question initiates an inquiry in to the nature of man, through a form of internal reflection; “I know will only by the sentiment of my own will, and understanding is no better known to me” (E 280). The will and conscience will concern us in section IV; for the present purposes, the discussion of understanding is most salient.

Rousseau’s inquiry into the understanding begins with reflection on its acts, on what it does, and not with something like an attempt to directly intuit its nature or to derive claims about that from metaphysical first principles. According to him, the primitive mental happening is sensing in which objects are presented “isolated, such as they are in nature.” Rousseau then tells us that he compares the materials of sensation: “I superimpose them on one another in order to pronounce on their difference or their likeness and generally on all

¹⁷ Is it irresponsible to thus identify the Vicar’s words with those of his author? In defense of doing so, note that Rousseau took the result of “the most ardent and sincere investigations ever conducted by any mortal” (R 54) to be “more or less what I have written down in my ‘Profession of Faith of a Savoyard Priest’” (R 55), while making perfectly clear that he is the one who executed these “unparalleled” investigations.

their relations” (E 270). Such comparing and pronouncing are really just judging (E 270). But judging and erring (which is simply the abuse of the capacity to judge) cannot according to Rousseau be accounted for by empiricist psychology as the effects of a “purely sensitive being:”

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 not force to bend them back on one another, it will never compare them. (E 270-1)

Not surprisingly, Rousseau goes on to claim that judging and misjudging are instead to be understood as the expression of an active force or power. The capacity to judge is “the distinctive faculty of the active or intelligent being,” since to have the power of judgment “is to be able to give a sense to the word *is*” (E 270); it is to have the capacity to think about how things are. Rousseau then identifies *himself* or “the individual *I*” (E 279) with this capacity:

Let this or that name be given to this force of my mind which brings together and compares my sensations; let it be called *attention, meditation, reflection*, or whatever one wishes. It is still true that it is in me and not in things, that it is I alone who produce it...I am not simply a sensitive and passive being but an active and intelligent being; and whatever philosophy may say about it, I shall dare pretend to the honor of thinking. (E 271-2)

After asserting that he “need only know that matter is extended and divisible in order to be sure that it cannot think” (E 279), he claims his active and passive parts are immaterial and material respectively. Striking as that claim might be, the heavy metaphysical reading it invites obscures the real import of Rousseau’s discussion. The important point is that in order for a certain kind of event to be an expression of an immaterial part, i.e., the soul, it must be *activity*, where this means an expression of “the power of comparing and judging” (E 280). Treating the soul as immaterial is simply a way of insisting on the primitiveness and irreducibility of the idea of activity and the active powers themselves; it is, I think, best not regarded as an insistence on substance dualism.

It is beyond doubt, Rousseau thinks, that empiricist psychology is unable to account for judgment and thought: whatever philosophy may say about these, it may *not* say that the purely sensitive being thinks.

It seems to me that far from saying that rocks think, modern philosophy has discovered, on the contrary, that men do not think. It no longer recognizes anything but sensitive being in nature, and the whole difference it finds between a man and a stone is that man is a sensitive being with sensations while a stone is a sensitive being without them. (E 279)

In so far as there is something left for philosophers to dispute, it is, perhaps, the claim that he thinks or that humans think. But if we grant Rousseau this starting point, he is confident that he has shown that we are active on the ground that such is a condition of the possibility of thought or judgment; “If I have just discovered successively these attributes of which I have no absolute idea, I have done so by compulsory inferences, by the good use of my reason” (E 286).

In the “Profession of Faith,” then, Rousseau distinguishes between his soul and body, his active and passive parts, his interior and exterior, according to whether such exhibit or are responsible for thought and judgment. The basis for the distinction here is of the utmost importance. We don’t make any progress towards understanding the soul by simple appeal to the concepts of activity, the inner, and immateriality. More generally, we don’t make any progress towards understanding any of these by appeal to any of the others. All these notions and the distinctions associated with them – soul/body, activity/passivity, inner/outer, immateriality/materiality – are subtly interconnected; each is susceptible to deployment and specification in widely various ways, and the shading given to any one will systematically affect the understanding of the others. That these are concepts which admit of several determinations is not unrecognized by Rousseau. Indeed, one could argue that his use of metaphors of biological defect (E 37, 254) and his discussion of animal instinct (E 286) indicate that he thinks that there is a sense in which plants, brutes, as well as humans, have souls, natures, or inner sources of change. In the context of Rousseau’s discussion of the understanding (which is the same context as his discussion of conscience), crucially, “soul” means *rational* soul, and “activity” means *rational* activity. When he identifies himself, or his real self, with the soul, his active nature, and what is in him, Rousseau’s attention is on what we might call intelligent activity, the intelligent soul, and reason as a source of change.

Now we are in a position to ask whether Taylor is entitled to hold that the motives arising from the soul are those that *really* originate within an agent, and that conscience is

such an inner source. One requirement on Taylor's interpretation is that it have a way to understand the connection Rousseau makes between morality and freedom. We have seen that Taylor's suggestion – one acts freely when one acts from motives which originate within oneself – will not do; “coming from within” isn't sufficiently fine grained to capture the special character of motivation by conscience. On Taylor's behalf, I suggested that we might distinguish those motives which *really* originate within an agent according to whether they are to be regarded as an expression of the soul. But the course of our discussion has shown that talk of the soul by itself will get us no farther than talk of the inner. When we pay attention to the primary context in which Rousseau discusses the soul, though, we can see that he has something more determinate in mind; namely, the rational soul. Given the way that judgment is linked with this primary use of “soul,” “activity,” and “inner,” it is not open to Taylor to say that conscience is part of the soul, while giving an account of it on which it is not an intelligent power – and on which it is merely a natural love of what reason anyway recognizes to be good, i.e., a mere passion.

IV. Conscience: “It is you who make the excellence of his nature and the morality of his actions.”
(E 290)

1. If Rousseau is to avoid the difficulties of rational intuitionism and if he is to hold that being motivated by one's conscience is what it is to be “really free,” then conscience must be something other than a passion sensitive to the judgments of reason and its particular expressions must be something other than mere feelings. His own appreciation of the active, rational character of conscience may be seen reflected in a key passage from *Emile*: “If one clearly understands that man is active in his judgments and that his understanding is only the power of comparing and judging, one will see that his freedom is only a similar power or one derived from the former. One chooses the good as he has judged the true.” Yet, Rousseau sometimes calls conscience and its particular expressions “sentiments,” and this has understandably led commentators to attribute a non-cognitivist theory of conscience to him. For someone aiming at such a reading, the following passages are the basic textual resources: “Reason alone teaches us to know the good and bad. Conscience, which makes us love the former and hate the latter, although independent of reason, cannot therefore be developed

without it” (E 67); “The acts of conscience are not judgments but sentiments” (E 290); “As soon as his reason makes him know [the good], his conscience leads him to love it. It is this sentiment which is innate” (E 290); “To know the good is not to love it; man does not have innate knowledge of it, but as soon as his reason makes him know it, his conscience leads him to love it” (E 290); “Did [God] not give me conscience for loving the good, reason for knowing it, and liberty for choosing it?” (E 294). Indeed, in his defense of a non-cognitivist interpretation of conscience as “a love of order” or “desire for the good” Jerome Schneewind cites most every one of these.¹⁸

In order to find our way into a proper understanding of conscience, then, while doing justice to Rousseau’s text, we must locate a sense of “sentiment” as something other than mere feeling, distinct from sensation, appetite, impulse and the like¹⁹ Against always hearing “sentiment” as “feeling,” consider the following passage from the “Profession of Faith:” “I am not propounding to you the sentiment of another or my own as a rule. I am offering it to you for examination” (E 260). An implication of this is that a sentiment *can* be commanded, even if in this instance, as it happens, the sentiment is not offered in that way. Similarly, Rousseau’s warning to “always remember that I am not teaching my sentiment; I am revealing it” (E 277), suggests that a sentiment is something that *can* be taught – what is clearly not the case with mere feeling. At the same time, a mere tendency or a brute inclination might be thought to be graspable in thought and teachable because it, unlike a particular feeling, appears to have a kind of generality and content. Supposing that this is true, we haven’t yet done anything to distinguish these from the hoped for sense of “sentiment.” Rousseau’s discussion of instinct proffers a potential key to this distinction.

For some, an instinct is simply an innate mere tendency to do *A* in *C*. Those suspicious that there is any such thing seem to be moved by a more general suspicion about innateness. They attempt to account for a pattern of behavior that is purportedly an expression of instinct by reconstructing the real origins of the habit: here an instinct is said to be merely “a

¹⁸ Schneewind 1998, 474-7.

¹⁹ In defense of such an interpretative gesture, we might note Rousseau’s own warning about the equivocal nature of his central theoretical terms: ; “I have a hundred times in writing made the reflection that it is impossible in a long work always to give the same meanings to the same words. There is no language rich enough to furnish as many terms, turns, and phrases as our ideas can have modifications” (E 108)

habit without reflection which is, however, acquired by reflecting” (E 287). But notice that on this understanding of instinct each of the following mere tendencies, if innate or unlearned, counts as one: (i) the tendency of zebras to drown when in rough seas, (ii) the tendency of lemmings to jump off cliffs when afraid, and (iii) the tendency of humans to drink when thirsty. However, it flies in the face of common sense to understand (i) or (ii) as instincts. But why? For something to be an instinct it has to have a point: as Rousseau tells us an instinct “appears without any acquired knowledge to guide animals toward some end” (E 287). But not just any end: instinct guides creatures to do those things that are good for it – which are suitable to its nature.

The nature of a thing is itself determined by whether it has what we might think of as a master instinct, what Rousseau calls an “original disposition:” “they are what I call in us *nature*. It is, then, to these original dispositions that everything must be related” (E 39). Sometimes he calls these “instincts,” sometimes “sentiments,” and sometimes “principles.” In addition, he calls their particular expressions “instincts” and “sentiments” as well. Of course, this can give rise to limitless confusion. To try to avoid that while motivating my cognitivist interpretation of conscience, I’ll examine the content of each of the three original dispositions described at E 39 with an eye to seeing how the sentiment, instinct, or principle of conscience is integrally bound up with the final one.

If we restrict our view to the merely animal world, self-preservation provides the standard according to which a pattern of behavior is determined to be natural (E 97). Of course, brutes do not have the concept of self-preservation. Still, nature has conveniently made it so that all of what living requires is pleasant to do, e.g., eating and sleeping; thereby a connection is formed between what is naturally good and what is pleasant. Indeed, Rousseau claims early on in *Emile* that our first nature – the first original disposition – is our being “disposed to seek or avoid the objects which produce sensations...according to whether they are pleasant or unpleasant to us” (E 39). The brute, like the child, doesn’t reason: “restricted to pure sensations” (DI 171) each has only “the sentiment of its present existence” (DI 151). However, unlike mere animal life, childhood is “reason’s sleep” (E 107). Waking up happens in stages, where each stage is governed by one of the forms of *rational* self-concern I considered earlier: *amour-de-soi* and *amour propre*. Where it is in the nature of animal life to

do what sustains it in doing what is pleasant, it is in the nature of animal life modified by reason to be moved by the *ideas* of self-preservation and pleasure. Our first second nature, as it were, is a self-conscious version of our original animal nature modified by technical abilities and a “mechanical prudence” (DI 171).

It is at this second stage that, strictly speaking, the life of the individual begins. It is then that he gains consciousness of himself. Memory extends the sentiment of identity to all the moments of his existence; he becomes truly one, the same, and consequently already capable of happiness or unhappiness. (E 78)

Here, reason is in the service of immediate and particular pleasures – say figuring out how to get to an apple which hangs from a branch twenty feet up – as well as more long term interests – say setting traps in summer for the autumn migration. Just as before, Rousseau articulates a standard or original disposition which governs this kind of life: “we are disposed to seek or avoid the objects which produce [sensations]...according to the conformity, or lack of it, that we find between us and these objects” (E 39), where our judgments of conformity are themselves governed by considerations of self-preservation and pleasure.

In a spirit similar to Kant’s well known teleological argument in the *Groundwork* (G 395-6), Rousseau suggests that it would be contrary to “the order of nature” were the purpose of reason the preservation and pleasure of the creature which has it; “if it [nature] destined us to be healthy then, I almost dare assert, the state of reflection is a state against Nature, and the man who meditates a depraved animal” (DI 145).²⁰ If reason is part of our nature at all, Rousseau would have us expect to learn of another an higher purpose. It is here that we are introduced to the third original disposition; it is the capacity to act “according to the judgments we make about them [sensations] on the basis of the idea of happiness or of perfection given us by reason” (E 39). Our second second nature, as it were, is the capacity to be moved by goods as good.

2. As Rousseau tells us, conscience “is to the soul what instinct is to the body” (E 286-7), and more specifically he tells us that it is “an innate principle of justice and virtue according

²⁰ These remarks are read rather differently by Cohen, who doesn't see that in this passage Rousseau is actually talking about health and not about something more grand like being well-ordered or being in accord with oneself. Cohen 1997, 117.

to which, in spite of our own maxims, we judge our actions and those of others as *good or bad*" (E 289, my italics). Conscience is the principle or measure by which we are moved by normative considerations, the possession of which is our ultimate original disposition. I propose to understand this principle in terms of its status as the principle of a being that acts "without anything external to me determining me," said with a very special emphasis (E 280). The fundamental point that will enable us to hear Rousseau's remarks on conscience in their proper register is this: *Rousseau thinks that acting and choosing are practical employments of the active power of reason*. Recall the passage quoted above: "if one clearly understands that man is active in his judgments, and that his understanding is only the power of comparing and judging, one will see that his freedom is only a similar power or one derived from the former. One chooses the good as he has judged the true" (E 380). In passages that purportedly support a non-cognitivist interpretation of conscience, e.g., "acts of conscience are not judgments but sentiments" (E 290), I take Rousseau to be emphasizing that acts of conscience are not theoretical judgments in which what is already the case is represented, but practical judgments in which what would be good is realized..

But what is the power of choice or freedom? In warning us not to mistake "unbridled license for freedom" (DI 120), I take Rousseau to mean that judging and choosing are judging and choosing according to a rule or law. Indeed Rousseau thinks that being free is having the capacity to make a judgment about what it is good to do "without anything external to me determining me" (E 280) and that he thinks this is the very same as the capacity to act and choose according to the moral law. We can find support for this reading in Rousseau's remarks about "pure spirits," what are, I take it, the equivalent of Kant's holy wills. Here's one: "Where our perishable needs end, where our senseless desires cease, our passions and our crimes ought also to cease. To what perversity would pure spirits be susceptible? Needing nothing, why would they be wicked?...they would be able to will only the good" (E 284). Of course, we humans aren't nearly this pure, and so in us the expression of this capacity takes the form of conscience; "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" (E 290).

Rousseau's conception of conscience is then very much like Kant's—"conscience is practical reason holding the human being's duty before him for his acquittal or condemnation in every case that comes under a law" (MM 399). For both Kant and Rousseau, conscience is "the condition of all duties as such" (MM 407).

Rousseau, like his great successor, finds a conception of conscience that eludes both the Scylla of Humean naturalism and the Charybdis of rational intuitionism by insisting upon the essential unity of rationality as it is expressed in cognitive judgment and in practical action. But our reading thus far leaves room for a question about whether conscience is actually "the work of the prejudices" (E 267), about whether morality is "a chimerical idea without any truth" (G 445). Rousseau has two responses to this kind of skepticism. First, he says that the skeptic doesn't prove that conscience doesn't exist, offers some empirical evidence that "all mankind" does as a matter of fact share a fundamental moral code, and emphasizes the strength of the testimony of his heart. But these points, especially the last, don't have quite the same force after Nietzsche's expression of the testimony of his heart, his *a priori*, and his discourse on the origin of conscience, as it were. So in the spirit of DI, I suggest that we set aside all the facts.

The second response, more compelling to us disenchanted moderns, is not to be found at any particular point in the text., But it can be reconstructed from Rousseau's explicit commitments through the whole of it. The argument, roughly, is as follows: Either we have *amour propre* or we do not. Suppose that we do not. Then, the whole range of social motives is not ours, something it would be hard for anyone to accept. So on the assumption that we do have *amour propre*, we must ask whether any substantive determination is internal to it. *Amour propre* inclines every animal that has it to rank itself in its own species (E 235, 278) and Rousseau thinks that this ordering can take one of two basic forms: "the good man orders himself in relation to the whole, and the wicked one orders the whole in relation to himself. The latter makes himself the center of all things; the former measures his radius and keeps to the circumference" (E 292). Rousseau argues that if our principle were inflamed *amour propre* (E 247), then we would be creatures who were destined to be unhappy – creatures which are "never content and never could be" (E 213). This is so because of the logic of competitive self-conceit: the sentiment of "preferring ourselves to others, also

demands others to prefer us to themselves, which is impossible” (E 213-4). ²¹In addition, on the hypothesis that *amour propre* is inflamed, instrumental irrationality would be ground into our nature because it “is forbidden to us by reason to want what we cannot obtain” (E 445). This intrinsic irrationality and our necessary unhappiness are each inconsistent with the hypothesis of the order of nature – a hypothesis that in some sense supplies the explanatory bedrock of *Emile*. Rousseau, after all, opens his “collection of reflections and observations” by claiming that “everything is good as it leaves the hands of the Author of things” (E 37). So, by elimination, conscience—the principle of the person who orders himself in relation to the whole—is the true shape of *amour-propre*.

V. A Final Problem

1. Thus far, we have seen that Rousseau rejects the empirical naturalist analysis of morality because it fails to account for the content and the form of moral thought and motivation. As a result, Rousseau accepts that the primary moral motive is the motive of duty – that recognition of the rightness of an action is the proper reason for doing it. The point is not simply that the moral worth of an action depends on the intention with which an action is done—Hume agrees with this—but rather that the only time an action has moral worth is when it is done because it is right.

But this is not yet to say anything about how it is that right action is to be identified or defined. The intuitionist agrees with Rousseau that it is only when an action is done because it is right that an agent deserves our esteem (RP 602). But Rousseau’s dismissal of intuitionist metaphysics brings with it the acceptance of a Humean style analysis of right action in terms of motives – an analysis of morality in terms of what goes on in human beings. In denying that the source of moral requirement is in nature, and affirming that it is in *our* nature in the form of conscience, Rousseau has in effect said that the way to understand what it is right to do is in terms of what conscience commands. This means that conscience must be able to independently identify which actions are right.

To hold these two theses, one about the proper motivation for right action and one about the proper analysis of which actions are right, is to do what Hume denied could be done. It is a rejection of Hume's dilemma and involves making a commitment to the centrality of the concept of self-determination or autonomy in one's moral theory. By rejecting the dilemma one marks out moral thought as a *wholly* independent sphere of practical thought – it is independent of lower sources of motivation and of the all too sublime intuitionist metaphysics²². So, Rousseau's conception of morality commits him to a picture on which there is no moral good identifiable independently of the commands of conscience. But, if conscience is to require us to do anything, then presumably Rousseau must be able to say how the standards of conscience specify particular actions as to be done.

Here is our problem: he does no such thing.

VI. The Quietism of Emile: "Do not expect lengthy precepts of morality from me. I have only one precept to give to you, and it comprehends all the others. Be a man." (E 445)

1. Rousseau offers no account of how the will is able to direct itself to action without depending on a principle which specifies the satisfaction of a motive as an end. This absence might be seen as decisive evidence against my interpretation of conscience; surely, Rousseau must hold a different view, one on which no such problem arises. But such a reaction, though understandable, is not obligatory. Instead, we ought to see the absence of such a discussion as a kind of quietism. Rousseau's silence is both philosophically and morally motivated. If this alternative reading of Rousseau's silence is tenable, it should serve as evidence, albeit indirect, for my interpretation of Rousseau's account of the nature of moral requirement. I will argue that the reason he halts his analysis of conscience where he does is that by his lights going further is ineffective and unnecessary in the course of an education, and so there is no reason we should expect to encounter such an extended account in a work such as *Emile*. More importantly, however, I'll argue that Rousseau thinks that pursuing such an analysis could be nothing other than an expression of vanity.

²² Note that this sense of independence is entirely separate from the conceptual and methodologically individualistic independence we saw as constraints on the empirical naturalist's account of motivation.

Rousseau thinks that providing a criterion for duties and a list of such duties would not effectively contribute to Emile's moral development. We can get the uselessness of explicating conscience in view by looking at Rousseau's attitudes towards methods of education in which one reasons with children (E 89). When considering instruction in theoretical matters, Rousseau says that "to substitute books for [experience] is not to teach us to reason. It is to teach us to use the reason of others. It is to teach us to believe much and never to know anything" (E 125). Belief in this sense is sufficient for impressing one's companions at dinner parties, but the context of its usefulness is limited to unhealthy forms of social dependence. For example, the truth about planetary movements is very little a part of what is involved in saying that "the earth goes around the sun" at a dinner party in order to make a good impression on others. It only matters that there is agreement that the statement is true. This kind of agreement, however, is irrelevant when predicting astronomical events. Such a distance from the object of one's beliefs indicates that this sort of education doesn't facilitate one's practical attachments to the world; moreover, as we'll see, Rousseau believes that it encourages a destructive separation from the world and one's freedom.

It should be unsurprising that the same criticism arises when Rousseau considers the method of teaching practical matters through rules or maxims; "put all the lessons of young people in actions rather than in speeches. Let them learn nothing in books which experience can teach them" (E 251). When direct experience isn't the most suitable means for moral education Rousseau recommends learning from detailed history or biography; but he is clear that this also does not involve setting out moral rules (E 248). Indeed, it is important that Rousseau thinks valuable moral learning can take place by reading history and biography, since it keeps alive the possibility that *Emile* and *Confessions* may be used in this capacity. Giving a set of rules to another merely as "the rules to follow" is to provide instructions without elucidating their goodness or point. The only incentive one has for following such rules grasped in this way is to receive the praise or avoid the punishment of the rule giver. In this sort of education, only an ability to imitate virtue develops but "all these virtues by imitation are the virtues of apes" (E 104).

For a moral education to be successful, the pupil must be taught to do what is right for the reason that it is right. In order for this to happen, the pupil must come to recognize the rightness of an action as a reason for doing it; only if we see the role of certain kinds of actions in the sphere of human purposes can we come to recognize doing what is right as good in-itself. “In doing good...one becomes good” (E 250), i.e., comes to do what’s good to do because it is good. This is because practicing virtue is naturally enjoyable and we most readily come to see this by experiencing this goodness for ourselves. Of course, Rousseau doesn’t think that everyone actually enjoys doing what is right, but just as “it is not for Slaves to reason about freedom” (DI 187), so I suspect he would say that it is not for the vicious or weak to reason about virtue. Since moral knowledge is practical, without addressing the motivational features that are part of such knowledge, one can only come to learn what others regard as required of us; “cold arguments can determine our opinions, but not our actions. They make us believe and not act. They demonstrate what must be thought not what must be done” (E 323). What is wanted from a moral education is to make knowledge of the practically necessary practically significant.

Given Rousseau’s views about the innateness of conscience and its implicit operations, specifying its content is unnecessary in the educational context. Let me begin with a nearly parallel case found in our theoretical lives. In order to teach a child to make good deductive inferences, one need not specify the rules that govern this activity. It is unnecessary, Rousseau would say because “reason is common to us” (E 266). I suspect that his claim would be that the situation is nearly the same with practical reason: goodness is common to us. However, just as it is necessary for a teacher to lead her pupil to make good inferences in order to successfully initiate the pupil into the rational order, so it is necessary for a teacher to lead her pupil to perform right acts in order to successfully initiate the pupil into the moral order. As Rousseau puts the point, “it is not by teaching the names of these virtues that one teaches them to children. It is by making the children taste them without knowing that they are” (E 131).

For Rousseau, a duty can’t be the kind of thing that must be learned from another. If he is right, then there is a moral argument for the claim that it is unnecessary to explicitly teach Emile what morality requires. In the “Profession of Faith” Rousseau says,

I shall never be able to conceive that what every man is obliged to know is confined to books, and that someone who does not have access to these books, or to those who understand them, is punished for an ignorance which is involuntary. Always books! What a mania... Either he will learn these duties by himself, or he is excused from knowing them. (E 303)

In modeling conscience on instinct, and developing a picture of the moral subject as autonomous, Rousseau restricts the role for moral education. It can't be a matter of telling a person what they couldn't know by themselves. Still, it might be a matter of getting a person to become sensitive to features of a situation to which she might not otherwise have been sensitive. If there is a problem about the accessibility of conscience, it is not a problem about the accessibility of its content, so not a problem that can be addressed by being told what conscience says; one still wouldn't have the ears for it. That is, moral education is best thought of as doing what is necessary to give the pupil a sensitivity to the promptings of conscience.

Furthermore, it would be morally wrong for the teacher to specify the rule governing the activity of conscience and the rules that conscience specifies as duties. As a result, Rousseau himself is constrained by conscience to abstain from such a thing. Before the age of reason, "appearing to preach virtue to children, one makes them love all the vices" (E 103). At this stage, when the moral sentiments and notions (E 219) are utterly foreign to one, being told "Do not lie" or "Be generous" encourages one to regard the demands of morality as interferences to one's happiness. It makes acting in accord with such commands unpleasant and encourages one to think that to violate those commands is to determine oneself. This is because at this early stage practical reason is guided by *amour de soi* and must see moral restrictions as mere obstacles to what one's will prescribes. Apart from making something other than a virtuous person, this kind of education engenders unhappiness because a pupil's conscience eventually demands (E 212) that he do exactly those things he has come to see as incompatible with his happiness and "self-determination."

2. But what about when Emile has reached the age of reason, when the moral sentiments and notions are not so foreign? This case is more interesting to a reader of *Emile* because the

reasons that Rousseau has for refraining from giving an analysis of conscience to Emile here will be reasons he has for refraining from addressing the reader in the same way.

We need to distinguish merely listing those actions specified as duties by conscience and giving an analysis of conscience itself. Simply to supply a list of duties claiming that the truth of the list is supported by philosophical argument without providing the argument would require the subjection of the recipient's reason to authority. The situation is the same when belief in revelation and miracles is required by the church: "a belief in all this on the faith of others, and a subjection of the authority of God, speaking to my reason, to the authority of men" (E 301). Providing a list would be providing a temptation, if not encouragement, to abandon one's responsibility for oneself; but "no one is exempt from the first duty of a man; no one has a right to rely on the judgment of others" (E 306). However, it seems that this difficulty can be resolved by supplying the missing argument, by showing how conscience specifies precisely those actions as duties. Besides, it is this analysis that is of real philosophical interest.

Here, I think, it is not silly to ask whether Emile might understand such complicated matters. Rousseau has doubts about the capacity of his own reason; "I knew, when I was pondering these things, that the human understanding, limited by the senses, could not fully comprehend them. I confined myself therefore to what was within my reach and did not attempt to understand what was beyond me" (R 59). Furthermore, the complexity would make it impossible to have the audience that he intends, namely, everyone; "O Man, whatever Land you may be from, whatever may be your opinions, listen" (DI 140). As long as Rousseau sees the sort of inquiry I've been suggesting others have found necessary to give as unnecessary for his own conduct, he will regard its results as nothing other than one of "the ills caused by our vain curiosity" (DAS 7). So we can see his silence on this matter, not as evidence for an alternative theoretical account of conscience, but as motivated by a duty to his reader and to himself; "I renounce idle questions which may agitate my *amour-propre* but are useless for my conduct and are beyond my reason" (E 277).

3. The points I've developed so far depend on construing *Emile* as a work with practical aims, in its own internal narrative and in its relation to its readership. But what about the

“Profession of Faith”, the most distinctively philosophical part of the book, a part set off from the rest in layout and narrative voice? Suppose that it is true that Rousseau intends *Emile* to affect the practical attitudes of his readers; isn’t it also true that parts of *Emile*, especially the “Profession of Faith,” are devoted to the unearthing of speculative truth? If the “Profession of Faith is to be seen as simply pursuing answers to a set of philosophical questions, then the above pragmatically “moral” considerations aren’t reason enough for keeping quiet in that context. So, one might argue, the absence of an analysis of conscience there is evidence either that Rousseau does not have the moral theory I’ve suggested, or that his account is a failure because he hasn’t shown how self-determination is possible. But this line of objection hinges on a misinterpretation of the role of the “Profession of Faith” in *Emile*. Rather than taking the “Profession of Faith” to provide the philosophical underpinnings of *Emile*, my suggestion is that we see it as an example of the kind of education one should give to a person who has already been corrupted by the influence of philosophy.

That *Emile* doesn’t hear the “Profession” shows that understanding it is not necessary for an education according to nature. Indeed, it is addressed to the reader with a directness unmatched in the body of *Emile*. The vicar says to Rousseau “if your sentiments were more stable, I would hesitate to expound mine to you. But in your present condition you will profit from thinking as I do,” to which Rousseau adds the following footnote: “this is, I believe, what the good vicar could say to the public at present” (E 295). But what is *our* present condition? In the *Reveries* Rousseau writes that instead of “removing my doubts and curing my uncertainties they [philosophers] had shaken all my most assured beliefs concerning the questions which were most important to me” (R 52). *Emile* confirms this; “in instructing him there about the religious controversy, they gave him doubts he had not had and taught him evils of which he had been ignorant. He heard new dogmas; he saw morals that were still newer to him. He saw them and almost became their victim” (E 260).

Rousseau is clear about the sort of address he wants the “Profession of Faith” to be both to its fictional recipient, who we later learn is Rousseau, and to the reader; “it is not my design here to enter into metaphysical discussions which are out of my reach and yours, and which, at bottom, lead to nothing. I have already told you that I wanted not to philosophize

with you but to help you consult your heart” (E 289). This is why Rousseau requests that we permit him “to leave aside Emile, whose pure and healthy heart can no longer serve as a rule for anyone, and to seek in myself an example that is more evident and closer to the morals of the reader” (E 344). If I’m right that the “Profession” is primarily addressed to the reader for the sake of dislodging morally dangerous philosophical prejudice, then the question of whether to include an analysis of conscience becomes a question about the degree to which such a discussion is necessary or sufficient for dislodging the relevant philosophical prejudices. It becomes a question of which philosophical thoughts are needed as “resources for living” (R 56).

For the most part, Rousseau takes writing philosophy to be what he is left with as a way of addressing the problem of modern man – that he is “born free and everywhere in chains” (SC 46). Describing his motivation to write *Emile* he says, “Not in a condition to fulfill the most useful task, I will dare at least to attempt the easier one; following the example of so many others, I shall put my hand not to the work but to the pen; and instead of doing what is necessary, I shall endeavor to say it” (E 50). Similarly, in the *Social Contract* he says, “if I were a prince or a legislator, I would not waste my time saying what has to be done. I would do it, or keep silent” (SC 46). However, philosophy itself becomes useful, and perhaps indispensable, when it is set to the task of correcting the corruptions of prior philosophical endeavors.

We see here, I think, a precursor of Kant’s thought that philosophical reflection is easily, frequently, and secretly put in the service of the unrestrained pursuit of the satisfaction of inclination, and a precursor of his own attempt to justify doing practical philosophy on the ground that it is the only way to counter these seductions (G 390, 405). However, while Kant thinks that when ordinary practical reason cultivates itself a natural dialectic ensues which “constrains it to seek help in philosophy” (G 405) and that it will find no rest “except in a complete critique of our reason” (G 405), Rousseau seems to think that the philosophical impulse is not a necessary or natural expression of reason in us, and that consequently something much less than a complete critique is required to squelch in someone the propensity “to put his hands over his ears and to argue with himself a little” (DI 162) for the sake of doing what he feels like doing. Rousseau can appeal to the difficulty of

the analysis of conscience as evidence for its frivolousness, while citing his own experience as evidence for the usefulness, and perhaps sufficiency, of the content of the “Profession of Faith”. Not only does Rousseau have moral reasons for not engaging in the sort of analysis Kant does, he makes no positive commitment to give such an analysis simply by having inquired about the nature of morality. This is because he sees philosophical inquiry as itself rooted in practical aims that wouldn’t be served by taking the “How possible?” question further than he has.

VII. Conclusion: “Others will perhaps demonstrate what I only indicate here.” (E 235)

1. I have argued for an interpretation of Rousseau’s moral theory on which the problem of content arises; at the same time, I claim that Rousseau has moral reasons for not spelling out its possibility. This allows us to read him as offering a theory of the will and an account of the nature of moral requirement that is more than an inspiration for Kant. According to Taylor, Rousseau’s role in the transition to Kant was as “the crucial hinge figure, because he provided the language, with an eloquence beyond compare, which could articulate this radical [Kantian] view.” But Taylor goes on to say the Rousseau didn’t think that “the moral law is what comes from within” and that it cannot “be defined by any external order.”²³ In contrast to Taylor, I take Rousseau to have done more than provide Kant with the language to talk about the will as autonomous. We can understand him as advocating a conception of moral philosophy grounded on a conception of the will as practical reason. And we can understand him as having such a view even though no action guiding principles of practical reason are specified, and even though conscience does not receive the title of “reason.” Once Rousseau’s thought is seen in this light, we might take the moves made by Kant regarding the question of the content of conscience as the development of an idea Rousseau thought needed no further discussion. It isn’t simply that there is no skeptical question to be asked, but that other considerations, namely practical ones, demand that he not inquire. Rousseau is not worried about discovering something horrible—that there is no principle, and so no

²³ Taylor 1989, 364.

morality. He is concerned about what he would have to be like to want to investigate such things.

Works by Hume

- T *A Treatise of Human Nature*. 2nd ed. Ed. L.A. Selby-Bigge. Rev. P.H. Nidditch. Oxford, 1978.
ECPM *An Enquiry Concerning the Principles of Morals*. Ed. J.B. Schneewind. Indianapolis, 1983.

Works by Kant

- G *Groundwork for the Metaphysics of Morals*. In *Practical Philosophy*. Ed. and transl. M.J. Gregor. Cambridge, 1996. Page references are to vol. 6 of the Akademie-Ausgabe, Kant 1902ff.
MM *The Metaphysics of Morals*. In *Practical Philosophy*. Ed. and transl. M.J. Gregor. Cambridge, 1996. Page references are to vol. 6 of the Akademie-Ausgabe, Kant 1902ff.

Works by Rousseau

- SC *On the Social Contract*. Ed. R. Masters. Transl. J. Masters. New York 1978.
E *Emile*. Ed. and transl. A. Bloom. New York 1979.
R *The Reveries of the Solitary Walker*. Transl. P. France. New York 1979.
DAS *The Discourse on the Arts and Sciences*. In *The First and Second Discourses together with the replies to critics and Essay on the Origin of Languages*. Ed. and transl. V. Gourevitch. New York 1986 (=Gourevitch (ed.)), pages.
DI *The Discourse on the Origin of Inequality*. In Gourevitch (ed.), pages.
EOL *Essay on the Origin of Languages*. In Gourevitch (ed.), pages.
RJJ *Rousseau, Judge of Jean-Jacques*. Transl. J. Bush/C. Kelly/R. Masters. Hanover 1990.

Other abbreviations

- L Hobbes, T. **Date**. *Leviathan*. In *The English Works of Thomas Hobbes*. Vol. III. Ed. W. Molesworth (reprint of the 1839-1845 edition). **Place of publication**. Cited by book, chapter, and section.
RP Price, R. 1991. *Review of the Principal Questions in Morals*. In *British Moralists*. Vol. II. Ed. D.D. Raphael. Indianapolis, 131-198.

- Cassirer, E. 1945. *Rousseau, Kant, Goethe: Two Essays*. Transl. J. Gutmann/P.O. Kristeller/J.H. Randall. Princeton.
—. 1951. *The Philosophy of the Enlightenment*. Transl. F. Koelin/J. Pettegrove. Princeton.
—. 1989. *The Question of J.J. Rousseau*. Transl. P. Gay. New Haven.
Cohen, J. 1997. "The Natural Goodness of Humanity". In *Reclaiming the History of Ethics: Essays in Honor of John Rawls*. Eds. B. Herman/C. Korsgaard/A. Reath. Cambridge, 102-39.
Kant, I. 1902ff. *Gesammelte Schriften*. Ed. Akademie der Wissenschaften (formerly Königliche Preussische Akademie der Wissenschaften). 29 vols. Berlin.
Korsgaard, C. 1989. "Kant's Analysis of Obligation: The Argument of *Foundations I*". *The Monist* 72, 311-340.
—. 1996. *Creating the Kingdom of Ends*. Cambridge.
Rawls, J. 1971. *A Theory of Justice*. Cambridge, MA.
—. 1993. *Political Liberalism*. New York.
Schneewind, J.B. 1998. *The Invention of Autonomy: A History of Modern Moral Philosophy*. Cambridge.
Taylor, C. 1989. *Sources of the Self*. Cambridge, MA.
Wood, A. 1996. "General Introduction". In Kant, I. *Practical Philosophy*. Ed. and transl. M.J. Gregor. Cambridge, xiii-xxxiii.