"From the Nature of Persons to the Structure of Morality"

Robert Noggle
Central Michigan University

Canadian Journal of Philosophy. Vol. 31, No. 4, December, 2001.

This is an Accepted Manuscript of an article published by Taylor & Francis in *Canadian Journal of Philosophy,* Vol. 31, No. 4, December, 2001, available online: http://www.tandfonline.com/

I. Intuitionism–in some form or another–is the most widely recognized and thoroughly discussed method of justification for moral theories. It rests on the claim that a moral theory must not deviate too much from our pre-theoretical moral convictions (or at least those that we are prepared to hold on reflection). In some form or another, this methodology goes back at least as far as Aristotle, and has been discussed, refined, and defended by such contemporary philosophers as John Rawls and Norman Daniels.[1]

---

[1]See John Rawls, A Theory of Justice (Cambridge: Harvard University Press, 1971), pages 46-53; and Norman Daniels, Justice and Justification: Reflective Equilibrium in Theory and Practice (Cambridge: Cambridge University Press, 1996). For a useful survey of recent literature on intuitionism, see Jeff McMahan, "Moral Intuition" in Hugh LaFollette, ed., The Blackwell Guide to Ethical Theory (Oxford: Blackwell, 2000). By "intuitionism" I mean the method of justification that came to be known by that name during the last 20 years or so, the method of which Rawls's notion of reflective equilibrium is a development; I am not, for instance, using the term "intuitionism" here to refer to anti-consequentialism or Ross-style pluralism (which is, ironically, how Rawls uses the term in A Theory of Justice [Cambridge: Harvard University Press, 1971]), or to restrict it to the very specific sort of program in moral epistemology that Sidgwick discusses in The Methods of Ethics (London: Macmillan, 1907). Instead, I simply mean the general thesis that moral theories must be checked against our pre-theoretical moral intuitions, or at least some subset of them, and the idea that those intuitions form one of the

There is, however, another methodology for constructing and defending moral theories. It draws on premises about human nature or the nature of persons to support conclusions about the nature and structure of morality. This method--which I will call the nature to morality methodology–evaluates a moral claim or moral theory on the basis of its relation to some (supposed) facts about the kind of beings we are. For brevity, I will use the term "nature-claims" to refer to claims about human nature or the nature of persons, and the term "nature-facts" to refer to true nature-claims. The nature-claims that have been used to support or criticize various moral theories include claims about human motivation, personal identity, the human soul, and the conceptual features of personhood or rational agency.[2] Like intuitionism, this methodology is as venerable as Aristotle, but as cutting-edge as the most recent debates about agent-centered

---

starting points for moral theorizing.

[2]For arguments that rest on claims about personal identity, see Derek Parfit, "Later Selves and Moral Principles," in Alan Montefiore, ed., Ethics and Personal Relations (Montreal: McGill-Queen's University Press, 1973), and part 3 of Reasons and Persons (Oxford: Oxford University Press, 1984). For arguments that rest on claims about the separateness of persons, see Rawls, A Theory of Justice, pages 23-4. Compare Robert Nozick, Anarchy, State, and Utopia,(New York: Basic Books, 1974) pages 32-3. For arguments that draw on claims about our motivational capacities, see Bernard Williams, "Persons, Character, and Morality," (in Moral Luck [Cambridge: Cambridge University Press, 1981]) and "A Critique of Utilitarianism" in J.J.C. Smart and B. Williams, eds., Utilitarianism: For and Against (Cambridge: Cambridge University Press, 1973) and Samuel Scheffler, The Rejection of Consequentialism (Oxford: Oxford University Press, 1982). Proponents of some versions of feminist ethics also base ethical claims or ethical theories on facts about human biology or psychology. See, for example, Carol Gilligan, In a Different Voice (Cambridge: Harvard University Press, 1982); Nel Noddings, Caring (Berkeley: University of California Press, 1984); Virginia Held, Feminist Morality (Chicago: University of Chicago Press, 1993), Sara Ruddick, Maternal Thinking (Boston: Beacon Press, 1989). See also Owen Flanagan, The Varieties of Moral Personality (Cambridge: Harvard University Press, 1991), pages 196-252.

prerogatives and restrictions. Unlike intuitionism, however, this method has received very little systematic attention.[3]

The idea underlying the nature to morality methodology is that the moral norms governing something must depend at least in part on what that something is.[4] If this approach to the justification of moral theories can be developed and defended, it may offer an alternative or at least a supplement to the intuitionist method. This paper will explore the logic, the presuppositions, and the defensibility of the nature to morality methodology. Because I want to grind only one ax at a time, I will investigate the methodology rather than using it to argue for a particular moral theory.

II. It was once common for philosophers to build and defend complete philosophical systems. Such a system usually included a theory of human nature and a theory of morality, and the more integrated the two, the better the system would hang together and the stronger its claim to be a complete and coherent unifying view of the world would be. When a moral theory is integrated with, derived from, or otherwise explicitly tied to a theory of human nature, it is easy to see how

---

[3]Some discussion appears in Samuel Scheffler, The Rejection of Consequentialism; Shelly Kagan, The Limits of Morality (Oxford: Oxford University Press, 1989) and Normative Ethics [Boulder, Colorado: Westview, 1998], pp. 280-94); Owen Flanagan, The Varieties of Moral Personality; Earl Conee, "On Seeking a Rationale" (Philosophy and Phenomenological Research 45 [1985]); Philip Montague, "The 'Negative' and 'Positive' Arguments of Moral Moderates" (Pacific Philosophical Quarterly 77 [1996]: 37-44); Christopher McMahon, "Expression Arguments in Ethics" (Pacific Philosophical Quarterly 69 [1988]: 325-341); John Rawls, "The Independence of Moral Theory" (Presidential Address, American Philosophical Association, 1974); Norman Daniels, "Moral Theory and the Plasticity of Persons" (Monist 62 [1979]); Robert Stern, "The Relation between Moral Theory and Metaphysics" (Proceedings of the Aristotelian Society 92[1992]: 143-59).

[4]Or, as Rawls succinctly puts it, that "the correct regulative principle for anything depends on the nature of that thing." A Theory of Justice, p. 29.

the credibility of the one can affect the credibility of the other. Thus, for example, Aristotle's moral theory loses some of its credibility when we abandon his theory of the soul, and Hobbes's moral theory becomes less plausible with the discrediting of Hobbes's oversimplified mechanistic egoism. These days, however, moral theories are seldom defended as parts of complete philosophical systems. Contemporary philosophers typically defend moral theories that lack (or have been reformulated to be independent of) any explicit connections with theories of human nature. For this reason, we might expect the nature to morality methodology to be largely irrelevant to discussions of contemporary moral theories. Yet this methodology is used quite frequently even in debates about contemporary "stand-alone" moral theories. And herein lies a puzzle.

Consider the debate over neo-Hobbesian moral theory. Many neo-Hobbesians advocate contractarian moral theory without asserting that we are solitary rational egoists who can only form a moral community by agreeing to a contract which protects our individual self-interests. Nevertheless, many critics dismiss neo-Hobbesian moral theory because of its alleged connection with what they see as a false egoistical picture of human nature. But if a neo-Hobbesian moral theory does not claim that we are solitary egoists, then how can the fact that we are not solitary egoists be an argument against it?

Similarly, most contemporary forms of consequentialism are formulated so as to remain largely neutral about human nature or the nature of persons. Yet critics have advanced influential objections against consequentialism that are based on the fact that persons have their own "personal points of view," which give them motivations that make consequentialist impartiality difficult if not impossible to achieve.[5] Others argue that consequentialism contradicts or neglects

[5]See Bernard Williams, "Persons, Character, and Morality," and "A Critique of Utilitarianism" and Samuel Scheffler, The Rejection of Consequentialism. A "personal point of view" is a ranking of states of affairs, connected with a specific person, which (typically) differs from an impersonal, objective, or impartial ranking.

the "separateness of persons"[6] because it treats persons as members of a collective "super-agent" (like cells in the body, or bees in a hive), trading off harms to one person for greater benefits to the collective. Again, we have a puzzle: If consequentialism makes no claims about human nature or the nature of persons then why aren't criticisms that rest on such claims simply irrelevant?

III. Nature-to-morality arguments of this sort seem to presuppose the following two claims. The first is that a moral theory can be "linked" to a nature-claim, even if it neither asserts nor implies that claim. (In contemporary discussions, the term "reflect" is sometimes used to point to such a linkage; however, such discussions rarely do more than hint at what this relation of "reflecting" is supposed to be.) Let us call a relation that holds between a moral theory and some nature claim a linking relation. So the first claim underlying the nature to morality methodology is that a moral theory can reflect or bear a linking relation toward nature claims even if it does not assert or imply them. The second claim that underlies the nature to morality methodology is that the plausibility of a moral theory can be affected by whether or not it reflects the truth about human nature and/or the nature of persons. This claim, which we can call a "linking principle," implies that the plausibility of a moral theory is at least partly a function of the links it bears to various nature-claims.

It is worth noting that while the nature to morality methodology may provide a potential alternative to intuitionism, nothing prevents us from using the two methods together. For example, a linking principle could be a component of a wide reflective equilibrium among moral principles, considered moral intuitions, and relevant background theories.[7]

---

[6]See Rawls, A Theory of Justice pp. 23-4; Nozick, op. cit.; and Dennis McKerlie, "Egalitarianism and the Separateness of Persons" (Canadian Journal of Philosophy, 18 [1988]: 205-26).

[7]See Norman Daniels, "Wide Reflective Equilibrium and Theory Acceptance in Ethics" (Journal

of Philosophy 1979, 256-82). A moral theory that bears a suitable linking relation to the nature-claims made by our best theories of human nature would, ipso facto, cohere with at least one background theory that seems relevant to morality.

However, the use of the nature to morality methodology need not imply any particular claim about how important (if at all) it is for a moral theory to cohere with our moral intuitions about right and wrong.

Nor does the use of the nature to morality methodology presuppose ethical naturalism, even though there is a sense in which it bridges the is-ought gap by linking moral prescriptions to descriptive nature-claims. The use of the nature to morality methodology does not commit one to a naturalistic meta-ethics because the linking principle on which it rests can be construed as having something like the following form: "If a being B has property P, then B is subject to moral norm N." Thus understood, the linking principle can be read as a conditional prescription. For the linking principle dictates which prescriptions can be commanded of beings of a given kind. Such principles can be seen as conditional prescriptions that make certain moral requirements conditional upon whether the being to whom they are directed has a nature of a certain sort. Because it is a very general, high-level conditional prescription, it can function as a kind of adequacy condition for other, lower level prescriptions–an adequacy condition that can itself be regarded as a prescriptive claim that applies to other prescriptive claims. The assertion that some moral prescription holds <u>conditionally</u> on the obtaining of some descriptive claim about the nature of the being to whom the prescription is addressed does not entail that the moral prescription is a descriptive claim. Nor does it entail that moral prescriptions are derivable from (or definable in terms of) <u>pure</u> descriptive claims.

IV. The nature to morality methodology rests on the thought that it is desirable for a moral theory to reflect the truth about human nature or the nature of persons. Stated this blandly, the claim is perhaps innocuous enough. But if it is to ground a method for assessing the relative credibility of competing moral theories, it must be spelled out far more carefully. One fairly plausible way to do so would be to claim that if a moral theory bears a linking relation toward a nature-claim, then that nature-claim should be true rather than false. Call this the <u>conditional</u>

claim. To make the conditional claim about some linking relation is not to demand that every moral theory must bear that linking relation toward the relevant nature-facts, nor is it to say that every moral theory that fails to do so suffers a loss to its overall credibility. Instead, it is to say that if a moral theory does reflect any nature-claims, then it loses credibility if those nature-claims turn out to be false. We can make this claim even if we would regard such links as being "optional" in the sense that a moral theory gets no "points off" for lacking them.

Conditional claims of this sort seem fairly plausible. Given that overall coherence is a theoretical virtue of conceptual schemes, it seems reasonable to say that a moral theory that coheres with the other parts of our conceptual scheme–namely the facts about human nature or the nature of persons–is preferable to one that clashes with them. It certainly seems odd for a moral theory to be linked to false nature claims, and it certainly seems better for it to be linked to true nature-claims. Conditional claims have the added benefit of not requiring or entailing any broader claim about whether all moral theories must reflect the nature-facts. In cases in which a moral theory seems to be linked to a false nature-claim, confining herself to conditional claims allows a critic to criticize one particular theory without having to defend a strong general thesis about the relation between morality and human nature and/or the nature of persons.

Certain recent uses of the nature to morality methodology seem to rest on conditional claims of exactly this sort. Some feminist and communitarian criticisms of contractarian moral (and political) theory can be plausibly construed as relying on the claim that contractarianism (or at least certain forms of it) is inadequate because it reflects a false picture of human nature as atomistic, pre-social, and self-interested.[8] While such arguments do not necessarily assume

---

[8]To take just to prominent examples, Michael Sandel can be plausibly read as making this kind of criticism against the contractarian liberalism of Rawls in Liberalism and the Limits of Justice (see especially page 14) and Virginia Held can be read as making this kind of criticism against contractarian (and other) kinds of moral theories in Feminist Morality: Transforming Culture, Society, and Politics (see esp. pages 57-63 and 193-214). For a fuller discussion of some classic

anything in general about whether a moral theory should reflect the truth about our natures, they do clearly assume that it is a mistake for a moral theory to reflect a false picture of human nature or the nature of persons.

The cogency of such arguments will depend, among other things, on how convincing it is to claim that the theory is indeed linked to the (allegedly) false picture of human nature or the nature of persons. In some cases, proponents of the theory under attack may escape the criticism simply by denying that the theory is best seen as an attempt to reflect any particular nature claims at all. And indeed this is just how some contractarians have responded to the communitarian/feminist argument just mentioned. Rawls, for example, has vehemently denied that his moral theory should be seen as resting on any "metaphysical" picture of the self, and David Gauthier has argued that the atomistic homo economicus of contractarian moral theory is not meant as a portrayal of how human beings actually are, but rather as a theoretical model that demonstrates how moral constraints can have a "non-tuistic" rationale.[9] The possibility of

---

communitarian literature that reconstructs its criticisms of contractarianism in this way, see Will Kymlicka, Contemporary Political Philosophy (Oxford: Oxford University Press, 1990), pages 207-15 For a fuller discussion of some classic feminist literature that reconstructs criticisms of contractarian (and other) approaches to morality in this way, see Rosemarie Tong, Feminine and Feminist Ethics, pages 49-63.

[9]See Rawls, "Justice as Fairness: Political Not Metaphysical" and see David Gauthier, Morals by Agreement (Oxford: Oxford University Press, 1986), pages 319-339, especially page 338. Such replies may involve showing that while the theory in question does make use of some model of human nature and/or the nature of persons, that model is not meant to embody any claim on the part of the theory about what persons are really like. Such claims will be most compelling when they are accompanied by an alternative explanation of the work that the model does within the theory. Thus Gauthier's remarks about the role of homo economicus in his theory suggest that the homo economicus model of human nature is meant as a counterfactual that is necessary to show

denying that one's theory is meant to reflect a claim about the actual nature of human beings or persons will vastly complicate any uses of the nature to morality methodology that rest on conditional claims. Since contemporary moral theories typically lack explicit ties or links to claims about human nature or the nature of persons, it will generally be plausible to defend a moral theory from criticism based on a conditional claim merely by arguing that the theory is not really mean to reflect any nature-claims at all, and that it therefore reflects no false nature-claims.[10]

This means that really fruitful uses of the nature to morality methodology are likely to require something stronger than the conditional claim. They are likely to need to claim that any moral theory loses credibility not only if it bears a linking relation toward a false nature-claim, but also if it fails to bear a given linking relation toward some true nature-claim. Call this the unconditional claim. To make this claim is to assert that any moral theory that fails to bear the specified linking relation toward the relevant nature-facts must pay a price in terms of its overall credibility. (How great a price is, of course, open to debate. Although this is an important issue, I won't pursue it here, and will deal only with the general claim that a moral theory suffers some

---

that moral constraints do not derive their rationality from social motivations, since they would be rational even if (contrary to fact), persons were asocial.

[10]Of course it may be possible to produce a compelling analysis of the moral theory under attack that makes it difficult to deny the link between it and some false nature-claim. However, such arguments are likely to face an uphill battle, since they will in effect be analyzing the "intent" of someone else's theory. That is, they will have to claim, more or less, that a theory is "intended" to capture or reflect some fact about human nature or the nature of persons. In general, it will be possible for the proponent of the theory to give an alternate account of the true intent of the theory, and in general when two people disagree about the intent of a theory, it is probably best to believe the person who holds that theory.

loss to its credibility if it fails to reflect the nature-facts.) Such a claim goes beyond discouraging a theory from clashing with the nature-facts; it demands that a theory cohere or "gel" with those facts by bearing some specified linking relation toward them. Because it demands that <u>every</u> moral theory reflect the relevant nature-facts in an appropriate way, it closes off the escape hatch that the conditional claim leaves open.

In the balance of the paper, I will consider three questions: What possible linking relations are there? Are there reasons that would support making the unconditional claim about any of them? And, if so, can any of them be used to help settle debates among competing moral theories? Although I will be investigating the methodology itself, it will be helpful to place the methodological analysis in the context of some particular debate within moral theory. The debate I will use is the debate between proponents of the kind of deontological theory that approximates common sense morality (and which thus typically includes agent-centered prerogatives and restrictions), and proponents of (act) consequentialist theories that require moral agents to bring about the impartially optimal state of affairs by any means necessary.[11] This particular debate is both important and one in which the nature to morality methodology has been often been used. As this debate has unfolded, it is generally the deontologist who seeks to use the nature to morality methodology to argue that consequentialism does not properly reflect human nature or the nature of persons.

---

[11]I shall use the terms "deontologist" and "consequentialist" to roughly correspond with Kagan's moderate and extremist. The deontologist, as I am thinking of her, denies that an agent is always morally required to (as Kagan puts it) "make her greatest possible contribution to the [impartially specified] greatest good." She may base this denial on an acceptance of agent-centered options and constraints, or on a skepticism about the existence of a single, rationally privileged impartial evaluative point of view that ranks states of affairs in such a way as to give all agents a moral reason to bring about the state of affairs that it ranks as optimal.

V. We have just seen that the nature to morality methodology will be most useful when it involves a linking relation of which the unconditional claim is plausible, that is, a relation that we can plausibly say that all moral theories should bear toward the relevant nature-facts. The most obvious candidate for a relation of that sort is the relation of logical consistency. After all, no acceptable theory may assert or imply false claims of any sort, and it is difficult to see why a moral theory should be any exception. So it seems reasonable to make the unconditional claim about the linking relation of logical consistency.[12] We might augment this claim by adopting Quine's principle that a theory implies the existence of any entity that it "quantifies over."[13] For example, a popular (though probably mistaken) reading of Kant's view claims that we exercise contra-causal freedom when we make moral choices. If this theory can be seen as making an existence claim on behalf of contra-causal freedom, and if it turns out that we do not have this kind of freedom, the theory would be in conflict with the nature-facts. Similarly, we could

---

[12]It turns out, however, that it would not be plausible to make the conditional claim about logical consistency. For making the conditional claim bout the relation of consistency would imply that a moral theory would become less plausible if it was consistent with any false nature-claims. This, implication, however, is absurd. Utilitarianism, for example, is not made less plausible simply because it is logically consistent with the false nature-claim that human beings are feathered tripeds. So we cannot infer the conditional claim from the unconditional claim: the fact that it would be plausible to make the unconditional claim about a given linking relation does not entail that it would also be plausible to make the conditional claim about that same linking relation. I thank an editor of The Canadian Journal of Philosophy for helping me to see this complication (and for avoiding a logical glitch in a previous way of formulating the ideas in this section of the paper).

[13]See W.V.O. Quine, "Existence and Quantification" in Ontological Relativity and Other Essays (New York: Columbia University Press, 1969), pp. 94ff.

imagine the following version of Derek Parfit's (pro-utilitarian) personal identity argument:[14] The (anti-utilitarian) moral doctrine that we may only trade off present harms to one person in order to secure benefits for a future version of that same person assumes the existence of the future version of the very same person. On Parfit's view of personal identity, this entity arguably does not exist. If he is right, then this moral doctrine is inconsistent with the facts about personal identity.[15]

Nevertheless, the requirement of logical consistency with the nature facts is likely be too weak to do very much work in terms of sorting out acceptable from unacceptable moral theories. For most moral theories will at least be logically consistent with the nature-facts, since (as we noted earlier) few contemporary moral theories make any explicit nature-claims at all. Those that can be seen as making explicit, clear, and unequivocal "existence claims" on behalf of non-existent entities seem likely to be fairly rare as well. Interesting and powerful uses of the nature to morality methodology are likely to require more robust linking relations than mere logical consistency.

Let us say that a moral theory recognizes a fact if that fact plays some definite role in the moral theory or its application. Recognition goes beyond mere consistency in that it requires some definite connection between the moral theory and the nature-facts. One very common and

---

[14]Parfit, op. cit. Parfit's actual argument is rather more subtle than the version given here.

[15] Yet another example of this argument was suggested to me by an editor for Canadian Journal of Philosophy. Virtue theory arguably assumes that there are relatively fixed character traits that can be the fundamental subjects of moral assessment. However, there is apparently a fair amount of psychological data that calls into question whether we do have anything like the kind of stable character traits that could qualify as virtues and vices in the way that virtue theory normally assumes. Thus it may be that the psychological facts are logically inconsistent with the key tenets of (at least some forms of) virtue theory.

very unproblematic way for a moral theory to recognize a nature-fact is for it to include subsidiary principles that tailor its application to those facts. For example, subsidiary moral principles requiring us to assist someone thrashing about in deep water allow a moral theory to recognize the fact that human beings lack gills. More generally, those moral theories that require us to benefit persons and to avoid harming them will normally include subsidiary principles about what sorts of things harm and benefit persons. I take it that this minimalistic sort of recognition can be required of all moral theories without further argument. To reject this claim is to allow that a theory can simply ignore information relevant to its application, and that seems to be decidedly irrational. But while this minimalistic kind of recognition is a slightly stronger notion of reflecting than mere logical consistency, it is still unlikely to discriminate acceptable from unacceptable theories. In particular, it is likely to be quite ineffective as a tool for the deontologist to use against the consequentialist. For consequentialist theories have little difficulty recognizing nature-facts in subsidiary principles that apply their more basic principle(s).

VI. Perhaps we can derive a more promising candidate for a linking relation from the Kantian "ought implies can" requirement that morality must only require what is possible. Thus, we might claim that a moral theory should reflect the facts about our natures by demanding of us only those things that are possible given the facts about the psychological make-up and the physical, intellectual, and motivational capacities of human beings. If this is a reasonable requirement, and if the nature-facts determine that some action, attitude, or way of life is impossible for us, then a moral theory that requires this action, attitude, or way of life fails to reflect the nature-facts and thus lose credibility.

Arguments of this sort are likely to appeal to the deontologist. Indeed, the claim that consequentialist theories require an impartiality and single-minded moral devotion that is impossible for us to achieve has been a common charge at least as far back as Mill. If morality

can require only what is possible, and if completely impartial benevolence is impossible, then any moral theory which requires it is in conflict with the nature facts–a conflict which costs it at least some of its initial credibility. Of course even if the empirical facts turn out the way the deontologist supposes, the argument will only go through if possibility is something that we can plausibly require of all moral theories (i.e., if we can defend the unconditional claim applied to the relation of possibility).

Now on the face of it, possibility requirements do have a fair amount of initial appeal. Moral philosophers as diverse as Kant, Aristotle, John Rawls, Owen Flanagan, and Stephen Darwall have argued that since morality provides guidance to human persons, it must be sensitive to their capacities. Indeed, the very notion of giving guidance--including moral guidance--seems to imply an ability to comply: Arguably one ceases to provide guidance if one recommends the impossible. And clearly the facts about human nature and the nature of persons do constrain what is possible for us. Thus we might think that those facts constrain what can be required of or recommended to us. Also, a morality that is too demanding might, as Flanagan says, "fail to grip us, and in failing to grip us . . . fail to gain our attention, respect, and effort."[16] A morality that requires the impossible may fail to motivate us to do even the possible, and thus may be practically self-defeating.

On the other hand, it is worth pointing out that a rejection of the "ought implies can" principle is not obviously crazy. Morality, one might think, demands whatever it demands, and if facts about our natures get in the way of our complying with its demands, then so much the worse for us. Although this Quixotic picture of morality may seem harsh or idealistic to many, it is not irrational or incoherent. The doctrine of original sin, after all, has had many believers, despite its implication that human nature inevitably makes it impossible to do what morality

---

[16]Owen Flanagan, Varieties of Moral Personality, p. 27. See also pages 15-55.

requires.[17]

Possibility requirements raise a further question: What notion of "possible" should we invoke? Possible for us as we are right now? Possible for us on a really good day? Possible for the best of us? Possible for us if we were fully to develop our natural empathy? Possible for us if we have enough therapy? Possible for us if we got rid of the psychological hang-ups that cause us to favor ourselves and our loved ones over strangers and to think of ourselves as separate individuals? A philosopher who claims that morality can only require what is possible for us must specify what exactly she means by "possible." However, even if she does this, it will remain open to her critics to suggest that she is not using the appropriate notion of possibility.

One way around this problem would be to specify the kinds of changes to our natures that a moral theory must not require. Thus we might hold that morality cannot require us to become the sort of tireless and fanatic Moral Saint that Susan Wolf describes, or to alienate ourselves from our most cherished projects and values in the way that Williams thinks that utilitarianism requires.[18] Thus the argument against utilitarianism would now take the following form: "While

---

[17]And one might reasonably distinguish between whether a theory is self-defeating and whether it is true. One might hold that if morality turns out to be so demanding as to be practically self-defeating, then that would be a reason to think that morality should only require us to promulgate a moral code that is as demanding as possible without causing people to give up on morality. Many people hold that an indirect or esoteric approach to morality like this would violate a so-called "publicity requirement" that rules out any moral theory that argues against publicizing its own actual demands. However, it seems fair to say that arguments in favor of such publicity requirements have been inconclusive at best.

[18]See Susan Wolf, "Moral Saints," The Journal of Philosophy 79 (1982): 419-439 and Bernard Williams op. cit. For the record, Williams also suggests that Kant's ethics is similarly alienating. I think that he is on somewhat less firm ground with this claim, for the jury is still out on the demandingness Kant's ethics. A good recent discussion of that issue is Marcia Baron's Kantian

it may not be, strictly speaking, impossible for us to behave as utilitarianism requires, we can comply with utilitarian demands only if we become very different from how we are now. This suggests that utilitarianism is unsuitable for beings like us."

While this further specification does eliminate much of the vagueness inherent in the claim that a moral theory must be possible in order to be plausible, it raises other problems. For notice that this version of the possibility requirement is only plausible if the change in our natures that would be necessary order for us to be able to obey the requirement in question would be a change for the worse. Thus the Wolf/Williams argument against utilitarianism is convincing only if becoming a utilitarian moral saint or being alienated from one's projects would be a morally undesirable change. But one might ask: "What is so good about being the kind of person with limited attachments, and what is so bad about being a moral saint?" Indeed, Shelly Kagan asks essentially this very question in reply to the kind of argument that Williams and Wolf offer. What Wolf and Williams see as a Good Thing, a crucial and valuable aspect of humanity, Kagan suggests is a Bad Thing, a selfish bias that morality should urge us to overcome.[19]

VII. This impasse over what conclusion to draw from the facts about our limited capacity for impartiality or benevolence reveals something important: What someone sees as the appropriate relation for a moral theory to bear to some nature-fact can depend upon her assumptions about the value of that fact. This should come as no great surprise, for in general, the appropriate response to a fact often does depend upon whether we regard the fact as a good thing or a bad thing. In particular, it is appropriate to celebrate and preserve a state of affairs that we regard as fortunate, and to lament or try to overcome a state of affairs that we regard as unfortunate. Thus, it would not be unreasonable to think that the relation that morality should bear to a nature-fact

---

Ethics (almost) without Apology (Ithaca: Cornell University Press, 1995).

[19]Kagan, pp. 271-78; 280-330, esp. 324.

might also depend on whether that fact is a Bad Thing to be lamented, or a Good Thing to be celebrated. Shelly Kagan makes the point this way: On the one hand, "we might picture morality as regretting the nature of persons, but realizing that it must recognize the facts for what they are, and asking what is the minimum number of concessions that must be made for them."[20] Or, we might see a certain nature-fact as something that "adds to the distinctive value and quality of our lives," and hold that "it is fitting for a moral system . . . to attempt to reflect the nature of persons fully--integrating, where feasible, features which are in harmony with" those facts.[21]

I will use the term "polarity" to mark the distinction between facts about–or features of–our nature that should be celebrated and those that should be regretted. A fact or feature worthy of celebrating has a "positive polarity," whereas one worthy of regret has a "negative polarity." It seems reasonable to expect morality to bear a very different kind of relation toward facts that have negative polarity than to facts that have positive polarity. Even if we agree that a plausible theory of morality should be linked to–and thus reflect–a given nature-fact, we might disagree about how morality should reflect it because we disagree about the fact's polarity.

---

[20]Ibid., 271.

[21]Ibid, p. 332. Rawls makes a similar point, noting that "moral conceptions regard persons differently and prize different aspects of their nature." ("The Independence of Moral Theory" p. 17.)

Non-consequentialists like Williams and Wolf may want to argue that the fact that humans have a personal point of view is a positive fact, something that should be celebrated. They may then claim that a credible moral theory should celebrate this fact by allowing persons to have an agent-centered prerogative to give extra weight to preferences that arise from their personal points of view. Consequentialists like Kagan, on the other hand, may want to argue that the existence of the personal point of view has negative polarity, that it is a selfish bias that is to be lamented rather than celebrated. They may concede that a moral theory should recognize the existence of the personal point of view (rather than merely ignoring it), but they may suggest that we regard it as a fact with negative polarity and try to overcome it or work around it in some way. Despite their differences, both parties agree that a moral theory should recognize the relevant nature-facts in some way. But they disagree about their assessment of the polarity of this particular nature-fact, and therefore about how morality ought to respond to it.[22]

The fairly reasonable idea that both parties to this debate seem to share is that morality should be structured around whatever is best about us, that it should be an extension or expression of, or a way of supporting, whatever is to be prized about our natures. In other words, we should want that those nature-facts that have positive polarity to have a strong effect on the basic structure of a moral theory. For example, we might expect that a plausible theory of morality should be structured so as to express, support, or extend such things as our ability to transcend our own point of view, our natural social impulses and empathy, or our rationality. On the other hand, a fact about us that is unfortunate and lamentable is not one that we should expect morality to express, support, or extend, though we may expect morality to take account of it in some way. Such facts should be recognized, but their effect on the structure and goals of the moral theory should be minimized. It is a fact that the moral theory should attempt to resist,

---

[22]I take my remarks here merely to summarize Kagan's remarks on pages 258-279 of The Limits of Morality, and to translate his point into my own terms.

counteract, or "work around" in some way. For example, we might expect that morality should counteract limited sympathies, that it should work around our natural partiality, and that it should resist our prejudices and baser impulses

VIII. Let us use the term "amelioration"for the relation that a moral theory bears to a fact when it treats it as an obstacle to be recognized but countered, resisted, or worked around. To ameliorate some fact is for the theory to recognize its existence but to treat it as an obstacle whose impact on the core structure of the theory is to be limited as much as possible. A theory that ameliorates some fact gives it the least possible effect on its basic structure while still taking it seriously rather than merely ignoring it. A moral theory ameliorates some nature-fact if it includes some feature designed to preserve the main idea or pursue the main goal of the theory in the face of some obstacle created by the nature-fact. For example, Kagan's theory about overcoming our natural selfish bias, or Mill's theory that moral education should at least partially harmonize personal interest with the collective interest would both count as ways to recognize the fact that human sympathies are rather limited. Both are attempts to work around or mitigate the fact rather than to give in to it by compromising the basic goals or structure of the consequentialist theory.

Let us use the term "embracing" for the relation that a moral theory bears toward a fact when it treats it as a fact to be celebrated and given an amplified influence on the theory's structure. To get a sense of what it might mean for a moral theory to embrace a nature-fact, consider the various ways that a theory could be said to embrace the separateness of persons or the personal point of view. A defender of a theory that includes agent-centered prerogatives to (sometimes) pursue one's own goals instead of bringing about the state of affairs most preferred from the impartial point of view could plausibly claim that her theory embraces the fact of the personal point of view. Similarly, a defender of a theory that includes agent-centered restrictions, or non-aggregation principles (such as those proposed by Taurek or Kamm) to

govern third party adjudication of competing interests, could plausibly claim that her theory embraces the separateness of persons.[23]

Intuitively, agent-centered prerogatives and restrictions, and non-aggregation principles seem to create a harmony between the moral theory and those facts. The problem, however, is that this notion of being in harmony, while intuitively easy enough to grasp, is not precise enough to decide among competing claims about how a moral theory might embrace a given nature-fact. For example, Samuel Scheffler imagines that a "sophisticated" consequentialist might argue that, despite appearances, consequentialism does embrace the fact of the personal point of view. Such a consequentialist might claim that morality should embrace the fact of the personal point of view not by including agent-centered prerogatives, but by making the aggregate satisfaction of preferences the sole criterion according to which we are to act.[24] She might argue that an aggregating consequentialism of this sort actually does embrace the personal point of view because the quantity that it seeks to maximize--aggregate preference satisfaction--is none other than the summing up of preferences arising from many personal points of view. Such consequentialist might offer this challenge: "If the personal point of view is so worth celebrating, then why not maximize the aggregate satisfaction of the preferences generated by every person's personal point of view?"

Thus we could have a dispute between defenders of competing moral theories over which theory really does embrace the fact of the personal point of view. A deontologist might claim that a moral theory can embrace the personal point of view only by including agent-centered prerogatives. Yet a consequentialist of the sort Scheffler imagines might claim that a moral

---

[23]See John Taurek, "Should the Numbers Count?" (Philosophy and Public Affairs 6 [1977]: 293-316) and Francis Kamm, Morality, Mortality, vol. 1 (Oxford: Oxford University Press, 1993)

[24]Scheffler, pp. 58-60; Scheffler uses different words, but I trust that I have presented the spirit of his idea. Cp. Kagan, op. cit., p. 345.

theory can embrace the personal point of view simply by requiring us to maximize the aggregate satisfaction of the preferences arising from every individual's personal point of view. Clearly, we will need to refine the notion of embracing if we are to settle such disputes.

Ideally, we will want a refinement of the notion of embracing to fulfill two desiderata. First, it should give us a reasonably clear way to tell which theories do embrace some fact and which do not. Second, we will want a refined notion of embracing that is clearly a way of "celebrating" a given fact. That is, given a refined notion of embracing, it should be clear why we should want a moral theory to bear this relation to those nature-facts that have a positive polarity.

IX. How then, might we refine the notion of embracing? What more specific linking relation might count as a way for a moral theory to embrace a given nature-fact? My first suggestion draws on the considerations about possibility that we discussed earlier (section VI). Let us say that a moral theory bears the <u>accessibility relation</u> toward some nature-fact if it does not require us to change that fact completely and permanently.[25] The accessibility relation satisfies the two desiderata that we just discussed. It seems to imply at least the outline of a decision procedure for determining which moral theories embrace a given fact and which do not. It also seems clearly to be an appropriate way for a moral theory to <u>celebrate</u> the fact in question. After all, if we think that some fact about ourselves is valuable in a way that morality should recognize, then it seems

---

[25]At least in "normal" circumstances. As an anonymous referee pointed out to me, a moral theory might call for us to give up what it sees as a good characteristic in very exceptional situations. I add the qualifications "permanently and completely" to deal with similar worries about moral theories that might <u>sometimes</u> ask us to temporarily subordinate one valuable aspect of our nature to some other value. Finally, it is important to note that a theory may be directly inaccessible by explicitly requiring us to change facts about ourselves, or it can be indirectly inaccessible by imposing requirements that can only be met if we change such facts.

reasonable to require that the moral theory not demand that we change those facts. Morality thus celebrates a feature of human nature or the nature of persons by not requiring us to give it up, and by claiming instead that retaining it is compatible with complying with our moral duties.

Armed with this understanding of embracing, the deontologist might argue that the fact that we are strongly motivated to be partial to our loved ones has a positive polarity and should therefore be embraced by any credible moral theory. Such an argument might motivate the positive assessment of the polarity of the fact of human partiality by appealing to considerations abut the nature of love, and about the fundamental significance of our most intimate bonds of love and affection. Such bonds, the deontologist might argue, are the highest flourishing of what is best about us, but (she might continue) they are by their very nature intimate and therefore partial. She might then argue that moral theories that include agent-centered prerogatives embrace the fact of our partiality by guaranteeing that we do not have to give up valuable features of human life in order to be moral.[26] Such prerogatives make morality accessible to us by guaranteeing that we do not have to completely and permanently give up our personal point of view in order to live a moral life. They allow us not only to keep our personal evaluative point of view, but to treat it as action-guiding. Thus they make morality accessible to beings like us by not forcing us to change the fact that much of our actions are guided by our own personal evaluative points of view.

---

[26]Of course more would need to be said about what counts as "giving up" a feature. Presumably the advocate of a partiality-respecting morality would want to allow that morality could, <u>in some situations</u>, require an agent to overcome or put aside her natural partiality. What she would want to rule out is a morality that made a permanent and pervasive move toward partiality a condition for being moral. She might also have to address indirect consequentialist arguments to the effect that some partiality may be an efficient means to promoting the impartial good. She may need to explain why retaining whatever degree of partiality is justified by these indirect arguments does not satisfy the condition that morality not require the giving up of one's natural partiality.

The accessibility relation seems to be a good candidate for a way to refine the notion of embracing. Let me suggest a second possibility. I will call it the "Rawlsian relation." In A Theory of Justice, Rawls sets out to study the relation between a moral theory and its underlying presuppositions by constructing and examining the "initial choice situation" from which it would be rational for hypothetical agents to adopt it. Rawls claims that his Two Principles of Justice would be chosen from an initial choice situation--the "Original Position"--in which the largely self-interested choosers know that they are (or will "become") one particular person living in the society for which they are choosing rules, but do not know which particular person they are.

By contrast, Rawls claims that it would be rational to choose utilitarianism from the following kind of initial choice situation: The chooser "imagines himself in the place of each person in turn, and when he has done this for everyone, the strength of his approval is determined by the balance of satisfactions to which he has sympathetically responded."[27] One would choose utilitarianism from this initial choice situation because one "chooses as if one will for certain live through the experiences of each individual, . . . and then sum up the result."[28] Rawls claims that the analysis of this initial choice situation shows that utilitarianism neglects the separateness of persons. For the chooser in the initial choice situation must be seen as living the lives of each person, and not of just one (separate) person.

The Original Position, on the other hand, models the impartiality of moral concern without neglecting the fact that human lives come "one to a customer." Thus, the separation of persons is built into the initial choice situation (the Original Position) that leads to Rawls's (deontological) Principles of Justice, but not the one that leads to utilitarianism. Similarly, the chooser in the Original Position knows that she has a personal point of view (although she is unaware of its contents), and she is concerned to choose principles that will allow her

---

[27] A Theory of Justice, page 186.

[28] Ibid, p. 189.

considerable freedom to pursue the projects it gives her, insulated to a large degree from any demand to promote the general welfare. Notice how the initial choice situation from which utilitarianism would be a rational choice involves constructing an agent (who bears a striking resemblance to the "ideal observer" of traditional utilitarian theory) who does not model the separateness of persons or the individual personal point of view. The choice is not made by individual agents, each of whom is merely one separate individual among many, and each of whom has a unique evaluative point of view. In this way, we give certain nature-claims a large effect on the structure of a moral theory by representing them in the initial choice situation from which a theory is to be chosen. The initial choice situation seems to link those nature-facts that it represents in its characterization of the choosing agents to the resulting moral theory.[29]

Readers familiar with debates about the Rawlsian project will be quick to note that the generation of a moral theory from an initial choice situation is neither simple nor

---

[29]However, there may be competing ways to model a given nature-fact. In The Possibility of Altruism (Oxford: Oxford University Press, 1970), Thomas Nagel suggests an initial choice situation in which the chooser will "branch" into each separate person who will eventually live in the society. He claims that a chooser in such a situation would endorse utilitarianism. Nagel's suggestion explicitly recognizes the separateness of persons. But unlike the chooser in Rawls's Original Position, Nagel's chooser does not become one of the persons but all of them. (Hare's argument for utilitarianism based on the [supposed] requirement to identify with each person would seem to be a development of this or a similar idea. See R.M. Hare, Freedom and Reason [Oxford: Oxford University Press, 1963], esp. chaps. 5 and 6.) While Nagel's proposed initial choice situation does ostensibly recognize the separateness of persons, it is unclear that a "branching" entity of this sort really is a good approximation of separate individuals. The branches of such an entity are less separate than actual persons. In Equality and Partiality (Oxford: Oxford University Press, 1991) Nagel seems to be suggesting that some version of egalitarianism unites an impartial moral concern with a respect for the separateness of persons.

uncontroversial.[30] I cannot address these worries here, but will simply say that if some form of the Rawlsian project of constructing initial choice situations that lead from a characterization of agents to distinct moral theories (or at least distinct families of theories) does turn out to be workable, then it may well provide a promising refinement of the notion of embracing for use in nature to morality arguments.[31]

A third candidate for a refinement of the notion of embracing emerges from the thought that the primary (though perhaps not exclusive) concern of morality is with the proper expression

---

[30]The most serious worry concerns the maximin decision rule which plays a crucial role in Rawls's derivation of a deontological system from the Original Position. Joshua Cohen's "Democratic Equality" (Ethics 99 [1989]: 727-51) is a nice discussion of these issues with references to the relevant literature.

[31]It is important to note that what I am calling the Rawlsian Relation is meant as a way to get a moral theory to embrace claims about human nature or the nature of persons, rather than a method for making a moral theory cohere with our considered moral intuitions about right and wrong. Although Rawls's own methodology emphasizes the task of getting moral theory to cohere with moral intuitions, it apparently does not rule out making it cohere with nature-facts. Indeed, Rawls seems to suggest (sec. 30 of Theory) that we can rule out the initial choice situation that leads to utilitarianism on the basis of objective facts about the separateness of persons. So it is possible to see the nature to morality methodology as an extension or elaboration of Rawls's own methodology. On the other hand, Rawls seems to disavow the nature to morality methodology in "The Independence of Moral Theory." I cannot hope to resolve this tension in Rawls's thought here, but I refer the interested reader to the superb discussion of this issue in Robert Stern, op. cit., pp. 146-151. As mentioned before, the nature to morality methodology is independent of any claim about how much stock we should put in a moral theory's ability to cohere with our moral intuitions.

of our concern for persons.[32] According to this line of thought, concern for persons is morally mandatory, and the role of a moral theory is to direct us in how best to manifest this concern. A moral theory will typically do this by articulating or at least presupposing a conception of persons which highlights features that are to be primary in our treatment of them. From this conception of the nature of persons, the moral theory will derive more specific requirements that articulate what a morally appropriate concern for persons amounts to. It will do this in part by specifying certain features of persons that are to be the primary focus of our concern for them.[33] On this view, the purpose of a moral theory is to tell us how to express our concern for persons by telling us what features are most significant about those persons for whom we are to be concerned.

For example, on one interpretation, utilitarianism can be thought of as regarding persons primarily as preference-satisfiers or experiencers. Utilitarianism seems to assume that we can adequately express our morally mandatory concern for persons (qua preference-satisfiers or experiencers) by attempting to bring it about that their preferences are satisfied or their experiences are pleasant.[34] Kantian ethics regards persons primarily as autonomous rational agents. Kantians claim that we adequately express our morally mandatory concern for persons

---

[32]In order to avoid repeating various qualifications, I will restrict myself here to that part of morality that concerns our treatment of persons. I do not mean to suggest that morality is only concerned with persons. In fact, the present line of thought is quite compatible with the idea that morality may be concerned with many other kinds of beings, and that its function is to show us the appropriate way to manifest our concern for each kind of being that has moral status.

[33]This is not to say that the further specification of what a person is fully determines what the moral theory will say; the moral theory will typically have to rely on other claims as well.

[34]Lawrence Haworth ("Autonomy and Utility" [*Ethics* 95 (Oct., 1984)]). and Will Kymlicka (op. cit., chapter 2) develop explanations of utilitarianism along these lines.

(qua autonomous agents) by respecting their decisions, refraining from treating them as mere means, and so on. In each case, a conception of persons highlights certain features as being most significant. The features toward which the theory directs our morally mandatory concern for persons dictate how that concern is to be expressed. Theories that do this appear, then, to be shaped in large part by nature-claims, and such shaping appears to constitute yet another linking relation. Let us say that a moral theory bears the transfer relation to the fact that persons have feature **F** if it derives requirements about how to treat persons from an implicit or explicit assumption that showing concern for persons qua bearers of **F** is tantamount to showing concern for persons themselves, or that we can adequately express our concern for persons by showing concern for them qua **F**-bearers.

The transfer relation appears to meet the second desideratum (see section VIII) for refining the notion of embracing. It seems reasonable that morality should require us to direct our concern for persons toward those features of persons which make them worthy of that concern in the first place–features that are to be regarded as valuable or which give persons their value or value-making status. The transfer relation does this by telling us to express our morally mandatory concern for persons primarily by expressing a concern for them qua bearers of those features. It allows those features of persons that are most valuable to shape the structure of the theory by guiding how the theory tells us to express our concern for persons.

Does the transfer relation meet the first desideratum? Can it tell us which theories embrace which claims about persons? Interestingly, Kantians do tend to emphasize the status of persons as unique, irreplaceable individual moral subjects. Focusing on these aspects of persons does seem to help motivate the Kantian idea that proper treatment of persons involves respecting the fact that they are inviolable as individuals who must not be sacrificed for the greater good.[35]

---

[35]There seem to be at least two different ways that Kantians try to move from this picture of persons to prohibitions against sacrifices. The first is to argue that because each person is an individual moral subject, she does not partake in a larger collective moral subjectivity. For that

Utilitarians, on the other hand, tend to emphasize the status of persons as experiencers and preference-satisfiers. Focusing on these aspects of persons does seem to help motivate the idea that proper treatment of persons involves ensuring that they have good rather than bad experiences, or that they have satisfied rather than unsatisfied preferences (or, perhaps, both). Thus we can see how differing conceptions of which features should be primary in our treatment of persons do seem to play a major role in helping to motivate different moral theories. This suggests that the transfer relation might well help us determine which moral theories embrace which features of persons.

In fact, a common Kantian argument against utilitarianism can be construed as being based on this line of thought. According to this argument, utilitarianism transfers moral concern to persons in an inadequate way, for it makes a serious mistake about the natures of persons by allowing them to be treated as mere means for the promotion of some desirable state of affairs. In so doing, it mistakes persons for mere objects which are suitable for use as tools to promote desirable outcomes, rather than recognizing the morally vital fact that persons are moral subjects and not mere objects. On the other hand, the utilitarian might argue that the appropriate way to treat persons is to advance their interests, and that therefore any policy that puts limits on advancing the aggregate interests of persons fails to treat them appropriately by failing to recognize by our treatment of them that they have interests, and that we most fully benefit persons in general when we promote their aggregate interests. The parties to this debate disagree

---

reason, benefits to other persons do not compensate her for the sacrifices she must make to achieve them. The second is to argue that the status of persons as autonomous moral subjects gives them a status that makes them unlike anything else in the world. And whereas most other things in the world are available for use as tools because they are mere objects, persons, because of their various properties (rationality, consciousness, autonomy, etc.) are not objects at all, but rather subjects. And while it is proper to treat objects as tools, it is not proper to treat subjects as tools, for they are of a different order of being, as it were.

about whether any fact about persons makes it inappropriate to benefit them simply by aggregating their interests and maximizing the extent to which this aggregate is promoted. Consequentialists tend to think that there is not, whereas deontologists tend to think that there is.

Notice that these arguments do not seem particularly compelling as arguments about which moral theories are logically implied by these facts about persons. Instead, they seem far more plausible as claims about a richer relation that a moral theory might bear toward a particular conception of persons. For while it seems rather dubious to claim that utilitarianism is logically inconsistent with the fact that persons are unique individual moral subjects, it does seem reasonable to suggest that utilitarianism directs moral concern primarily toward the status of persons as loci of preferences and experiences rather than to their status as unique individual moral subjects.[36] Similarly, while it seems dubious to claim that Kantian ethics is inconsistent with the fact that persons are experiencers and preference-satisfiers, it does seem reasonable to suggest that Kantian ethics advocates a transfer of moral concern to persons conceived in such a way as to emphasize their status as uniquely individual moral subjects.

Now of course it seems most plausible to claim that persons are both loci of preferences and experiences and uniquely individual moral subjects. So the main difference between the conceptions of personhood underlying utilitarianism and Kantian ethics may be one of relative

---

[36]Not all utilitarians see their theory in these terms, however. I discuss the relation of utilitarianism to conceptions of what is valuable about persons in "On the Cross of Mere Utility: Utilitarianism, Sacrifices, and the Value of Persons" (Utilitas 12 [2000]: 1-24). In that paper, I note that some utilitarians may accept what may be thought of as a somewhat more Kantian conception of persons. Part of my argument in that paper is that if a utilitarian does accept a more Kantian picture of persons (and in particular, the relation of persons to value), then utilitarianism may not turn out to have all of the (generally counter-intuitive) implications that have often been attributed to it. I suspect that the Kymlicka/Haworth construal of utilitarianism would fit this model.

emphasis. One way of construing this difference is to think of it as a difference about which features of persons are most central to their value, that is, which features have positive polarity (or the greatest positive polarity). Utilitarianism and Kantian ethics can be seen as attributing (the greatest) positive polarity to different aspects of persons, and directing our moral concern accordingly. Thus we can see both theories as being based on a transfer of concern to what they take to be the most important features of persons, and as disagreeing about what those features are.

X. Having three notions of embracing might seem an embarrassment of riches. For we wanted a relatively clear way of refining the notion of embracing, and instead of one, we have three. Fortunately, however, I think that all three of the proposed refinements of the notion of embracing can be reconciled within a single complex embracing relation.

We might combine the accessibility relation and the Rawlsian relation (assuming that it can be specified in such a way as to be useful) by extending the notion of accessability to include a kind of "voluntaristic accessibility." The idea here is that a theory that it would be irrational for an agent to choose (under suitably specified conditions) is inaccessible to her in the sense that it lies outside her range of rational choices given the kind of being she is. Thus we might think of the Rawlsian relation as a fuller kind accessibility than mere possibility: The fact that a being could choose a given moral theory in a situation in which she has and values certain facts about herself might be thought to show that the theory is accessible to her in the sense of being more fully compatible with the kind of being she is. Thus we might expand the notion of accessability to a given kind of agent to include a theory's choosability to that agent from an initial choice situation in which the most important features of that agent are represented.

In addition, I think that the transfer relation and the accessibility relations can also be seen as complements rather than competitors. For accessibility and the transfer relation pertain to very different aspects of persons. Accessability pertains to persons as moral agents. It is a

relation between morality and what persons may do. The transfer relation pertains to persons as moral patients. It is a relation between morality and what may be done to persons. We can thus define a complex embracing relation as the relation that holds between a moral theory and the facts about human nature and/or the nature of persons when the moral theory's treatment of persons as agents bears the accessibility relation (and perhaps the Rawlsian relation or some suitably revised variant) to the nature-facts, and when its treatment of persons as moral patients bears the transfer relation to the nature-facts. From now on, when I speak of embracing, I will mean this complex relation.

XI. Those who wish to argue that a moral theory should embrace those nature-facts that have positive polarity (typically it will be the deontologist making such arguments) face the challenge of defending their assessment of the polarity of those facts that they think a moral theory should embrace. For such arguments rely on premises that go beyond mere descriptive facts about human nature or the nature of persons to include evaluations of those facts. This complicates matters considerably, since a person's evaluation of a given nature-fact will often be part and parcel of the very same moral intuitions that push her toward certain moral theories and away from others. This means that people who hold different moral theories will often also hold different assessments of the polarity of the relevant nature-facts. Thus, for example, a negative attitude toward the personal point of view (a tendency to see it as a selfish bias) is likely to be part and parcel of the same intuitions that lead someone to defend consequentialism, while a positive attitude toward the existence of the personal point of view is likely to be part and parcel of the same intuitions that lead someone to favor some sort of deontological moral theory. In fact, many of our intuitions about the polarity of certain nature-facts are probably results of our accepting a moral theory, instead of independent reasons for it.[37] This means that intuitive

[37]I thank an Editor for Canadian Journal of Philosophy for this point.

assessments of polarity are unlikely to provide either an independent basis on which to build a moral theory or common ground from which to adjudicate debates between existing theories.

Despite these worries, however, there may be ways to convince one's opponents to change their evaluation of the polarity of the fact in question. Because such an evaluation is likely to be a fairly basic intuition, this undertaking might look less like an argument and more like an invitation to alter one's attitude toward the fact. But there might be reasonable ways to issue and assess such invitations. For example, suppose we could show that morality must rely in some way on a certain feature of persons. Such a feature would seem to be a strong candidate for having a positive polarity–a polarity "assigned" by morality itself. What features of persons are plausible candidates for being the "allies" of morality? Presumably those that morality "puts to work" in some crucial way. Presumably rationality and at least some form of autonomy are features of this sort.[38] Some might argue that the property of being an agent extending through time is also such a feature.[39] Arguably the capacity for the kind of critical reflection needed to distinguish between values and mere wants is also a necessary "ally" of morality. Suppose that we could show that morality must build upon and develop some feature of persons, or that a certain feature of persons is necessary for them to comply with moral demands. There would seem to be some reason to take these features to have a positive polarity. It would seem odd for morality to fail to assign positive polarity to features necessary for the kind of agency, actions, or life that morality itself commands.

---

[38]One could draw on work of Kant and Alan Gewirth for arguments along these lines. Gewirth has arguments that take this form, although they are embedded in a very different overall method of justification. See Reason and Morality (Chicago: University of Chicago Press, 1978).

[39]An argument of roughly this sort appears in Richard Lippke, "Why Persons Are the Ground of Rights (and Utility Isn't)" (Journal of Value Inquiry, 18 [1984]: 207-217, 209ff.).

XII. However, there may be a sounder footing for the claim that morality should embrace certain facts about human nature or the nature of persons. We might claim that morality must embrace certain nature-facts because they are metaphysically or conceptually fundamental. It seems plausible to think that if persons play a central role in moral theories, then those facts about them that are metaphysically and conceptually fundamental should be given a role in shaping the structure of any plausible theory about morality. In fact, if we think that morality ought to reflect any facts about the nature of persons at all, it seems reasonable to think that it must, first and foremost, reflect those facts that are definitive of what a person (or, alternately, what a human being) is.[40]

I think that there is a good case to be made for thinking that the complex embracing relation is an appropriate linking relation for a moral theory to bear toward metaphysically and conceptually fundamental facts about the kinds of beings we are. Consider each component of that relation. It makes sense to require moral demands to be accessible to us without our having to alter any facts about us that are deep metaphysical or conceptual truths about what it is to be an agent or a person. Indeed, this seems to be an extremely plausible possibility requirement, for it implies that any moral theory that would require a person to cease to be a person at all would be implausible as an account of the moral obligations of persons. This claim can be supported by noting that it seems to be part of the concepts of personhood (or the closely related concept of agency) and of morality that persons are subject to moral requirements but that non-person are not. Morality is, in an important sense, addressed to persons in a way that it is not addressed to non-persons. While we can use evaluative language in talking about events that involve non-

---

[40]I suspect that arguments relying on premises about fundamental features of persons will be more convincing than those relying only on premises about features that are fundamental to human beings, but which could easily be imagined to be absent in non-human persons, but I won't pursue this issue further here.

persons, non-persons cannot be blamed or praised on the basis of whether they conform to moral requirements. We may label a hurricane a Bad Thing, but we would not hold it morally blameworthy for the havoc it wreaks. Persons are subject to moral requirements in a way that non-persons cannot be. If being a person is a necessary condition for being subject to moral rules, then it seems reasonable to require that a morality whose point is to guide persons must not require that they cease to be persons. Moral requirements, it seems, must be such that it is possible for persons to obey them without ceasing to be persons.

Of course one might wonder whether this requirement is thick enough to do any significant work. That will depend on what exactly the metaphysically and conceptually fundamental facts about persons turn out to be. We should not assume prematurely that the concept of a person is too thin to do any significant work until we are sure we know what a fully developed conceptual and metaphysical analysis of personhood would look like. For example, some communitarian and feminist work seems to suggest (not implausibly) that a fully worked out conceptual and metaphysical analysis of personhood might be thick enough to place significant constraints on moral (and political) theories. Some of this work suggests that a full analysis of personhood may show that persons are, as a metaphysical or conceptual matter, individual moral subjects with unique evaluative perspectives which will generally diverge from any impartial evaluative perspective. Such work may eventually lay the foundation for a convincing argument that certain kinds of partiality are conceptually or metaphysically essential to personhood (or at least social personhood or personhood which is biologically instantiated). Of course there is far more to be said about such questions, but for present purposes, it is enough simply to suggest that the requirement that moral theories must not demand that persons cease to be persons may turn out to be a rather substantial constraint after all.

The transfer relation (which is the "moral patient" aspect of the complex embracing relation) also seems to be a plausible relation to require a moral theory to bear toward metaphysically and conceptually fundamental facts about persons. For if part of the point of

morality is to secure the proper treatment of persons by other persons, then it would appear that it must be constrained by the metaphysically and conceptually fundamental facts about persons. For it seems plausible to think that the proper treatment of a thing would depend on the metaphysically and conceptually fundamental features that are essential to or constitutive of that thing's being what it is. The transfer relation can be seen simply as a "substitution principle" that allows us to replace the generic term "persons" with a more specific description of that toward which we are to direct our moral concern.

Both parts of the complex embracing relation seem, then, to be appropriate relations for moral theories to bear toward the metaphysically and conceptually fundamental facts about persons. This use of the notion of embracing requires no evaluation of polarity, but it does require an assessment of the metaphysical or conceptual fundamentalness of nature-facts. If this can be given, then we may get round the problem of circularity of polarity arguments. Of course, it is an open question whether any metaphysically or conceptually fundamental nature-facts are sufficiently interesting to provide enough leverage for the complex embracing relation to favor certain moral theories over others.

Notice, however, that it will make a great deal of difference whether an argument is based on the claim that morality should embrace a given fact about us because of its value (polarity) or on the claim that it should embrace a fact because it is a metaphysically or conceptually fundamental part of what it is to be a person. For in the first case, we will need to make and defend an evaluative claim about the nature-facts that we want morality to embrace. In the latter case, we will not.

Thus, it will be important to keep track of exactly what is being claimed of a particular nature-fact. For certain nature-facts may be metaphysically and conceptually fundamental, while others may have a positive polarity; still others may be both fundamental and positive (indeed, some may be positive <u>because</u> they are fundamental). For example, the fact that humans seem prone to form emotional attachments to those with whom they reproduce seems to be a deep fact

about human nature, and surely one that is extremely valuable. Yet it may not be conceptually or metaphysically fundamental. If not, then arguments claiming that morality should embrace those facts will require yet another step–one that would support a positive assessment of the polarity of the facts. Given the difficulty of defending claims about the polarity of nature-facts, the claim that moral theories should embrace metaphysically and conceptually fundamental nature-facts may turn out to be more useful than the claim that moral theories should embrace positive nature-facts.

XIII. At this point we can draw some tentative conclusions about which theories seem to embrace which facts. If we refine the notion of embracing in the way that I have suggested, then it seems that a typical deontological theory will do a better job of embracing the separateness of persons and the personal point of view than will a typical consequentialist theory.

For we have seen that the agent-centered prerogatives that deontologists typically favor not only recognize the existence of the personal point of view, but they give each agent the permission to be guided by it (within certain morally-defined limits, of course), instead of requiring her to be guided always by the impartial point of view. It is not just that the interests generated by the personal point of view are valued by theories with agent–centered prerogatives; as we saw, the consequentialist can also argue that a theory which aggregates concerns arising from each person's point of view also values those interests. What is different is that the deontological theory does not require the agent to give up the personal point of view as her own unique evaluative point of view which is a direct source of reasons for her. The agent-centered prerogative allows the agent to retain a direct relationship to the interests that arise from her own evaluative point of view, and to allow those interests to play a direct (though not absolute) role in guiding the agent's own actions and life.

While (act) consequentialism can aggregate the interests that arise from the personal point of view, it does not allow the agent to keep her point of view as a distinct, action-guiding

evaluative framework. For this reason, consequentialism does not embrace the personal point of view to the same extent as deontology; at best consequentialism can only embrace the fact that the personal point of view is what generates (at least some of) the interests that it aggregates and maximizes, whereas deontology embraces the personal point of view from which those interests arise in the first place. The accessibility relation helps us to see this. Consequentialism allows us to retain our interests, but it aggregates them with everyone else's interests. What consequentialism does not allow us to retain (but which deontological theories do) is our own unique evaluative point of view.

Theories that include agent-centered-restrictions bear the transfer relation to the separateness of persons in a way that aggregate-maximizing consequentialist theories do not. To see this, consider how the plausibility of an aggregate-maximizing consequentialism like utilitarianism would change if we applied it to beings who really were best thought of as members of a group entity. That is, imagine a kind of being whose primary moral significance is not "located" at the level of the individual organism. The cells of a human body, or perhaps bees in a hive might qualify. For such beings, the primary moral significance is located in the whole collective rather than the individuals that constitute it. For this reason, harms to one individual can be unproblematically offset by greater benefits to the collective, for the benefits accrue to the being that is the primary locus of moral significance while the harms occur to the being whose moral significance is merely derivative.

On the other hand, a being who has its primary moral significance as an individual cannot be compensated for harms to itself just by having a greater benefit bestowed on the collection of other beings. Because it is a locus of moral significance separate from and prior to that of the collection, to benefit the collection is not necessarily to benefit the individual. Such an individual does not "partake" of the moral significance of the collective (which gets the benefit) in the way that a member of a collective entity does, so the benefit to the collective (or to other individual members) does not offset the harm to the individual in the way that it would if that individual

lacked an independent moral significance.

The fact that aggregation is seems more appropriate for beings whose primary locus of moral significance is in the collective than for beings whose primary locus of moral concern is in the individual does not, by itself, prove that aggregation is inappropriate for beings who have their moral significance primarily as individuals. Instead, it merely suggests that an aggregate-maximization theory does not transfer moral concern to the separateness of persons. This is not to say that it is logically inconsistent with the separateness of persons. Rather, it is only to say that aggregate-maximizing theories do not embrace this separateness; they do not reflect it in the same robust way that a theory with agent-centered restrictions does. If this is correct, then it is only the beginning of the argument, not the end. Even if we grant the principle that a plausible moral theory must embrace the most significant nature-facts, there still remains the question of whether the separateness of persons significant fact (assuming that it is a fact) about persons; that is, whether it is either strongly positive in polarity or metaphysically or conceptually fundamental.[41]

XIV. The underlying intuition behind the claim that morality ought to bear a robust relation such as embracing to human nature and/or the nature of persons seems to derive from a certain conception of the subject-matter and point of morality. According to that conception, a fundamental (and perhaps the fundamental) concern of morality is to secure the proper treatment of persons by other persons. If we assume that the structure of a theory must be shaped by its subject matter, then it stands to reason that a moral theory must be shaped by the facts about

[41]The consequentialist could claim that separateness is neither a positive fact nor a metaphysically or conceptually fundamental fact about persons. Taking this line would not necessarily involve rejecting the claim that morality should embrace the most significant facts about persons. Instead, it would simply deny that separateness is one of the most significant facts about persons.

persons. The principle that a plausible moral theory must bear the complex embracing relation to the facts about persons gives substance to this line of thought by specifying the way that the nature of persons can shape the structure of a moral theory. The fundamental intuition behind it is that morality should call for treatment of persons that responds to whatever is most significant about them, and that it should make demands on agents that are compatible with their retaining whatever features are most significant about them. If morality is largely person-centered, then it seems reasonable to expect a moral theory (or at least that part of it that deals with persons) to embrace the most important features of persons. This is not, of course, a knock-down argument, but I do think that it should give considerable comfort to those who pursue arguments that rest on this principle.

Consider what would be involved in its denial. Suppose, for instance, that a nature to morality argument could be put forward to show that consequentialism does not embrace certain important facts about human nature and/or the nature of persons. (The facts might be important in the sense of having a positive polarity, or they might be metaphysically or conceptually fundamental, or both.) Now one very direct consequentialist reply would be simply to reject the conception of morality as being fundamentally concerned with persons. Without that conception of morality, the claim that morality must embrace what is most significant about persons seems far less compelling.

Now some consequentialists may be prepared to do exactly this. They may <u>want</u> to do this because, as we have just seen, it appears that the version of the nature to morality methodology that involves the requirement that moral theories should embrace the nature-facts is likely to be most helpful to the deontologist. And they may be <u>able</u> to do this because their own theory gives them a way to articulate a competing conception of what morality is all about. Thus some consequentialists may claim that the concern of morality is not with persons <u>per se</u>, but with <u>states of affairs</u> and the <u>value</u> they contain. This conception of morality seems to offer little rationale for the requirement that a plausible moral theory must embrace the most significant

facts about persons.

Notice, however, that the consequentialist can thoroughly reject the conception of morality as being about persons only if she denies that we should care about the value of states of affairs because we care about persons (and perhaps other sentient beings) who partake in or create that value. For if we claim that a state of affairs is good simply because it is good for or valued by persons, then to care about producing good states of affairs is just to care about benefitting persons, and thus we have accepted the conception of morality that grounds the nature to morality methodology after all. Thus, it seems that if the consequentialist wants to immunize her theory completely against the most robust forms of the nature to morality methodology, she must claim that morality is about producing good states of affairs, and that the goodness of states of affairs does not derive from their goodness for or to persons.[42]

Certainly some people have made claims of this sort--Plato and G.E. Moore come to mind. The question, though, is whether the price of such immunization is worth paying. Taking value to be a fundamental concern of morality will, of course, sound right to many consequentialists. However, the view that value is fundamentally independent of persons is likely to seem at best unmotivated and at worst counter-intuitive, even to many consequentialists. Indeed, for many people, one of the main appeals of consequentialism is its ability to express a prior and more fundamental concern for persons.[43] Some consequentialists may be willing to pay

[42]I thank an editor for the Canadian Journal of Philosophy for pressing this point.

[43]Indeed, sometimes consequentialists will be pushed to this view of the underlying rationale for consequentialism when they are faced with arguments about how consequentialism neglects persons in favor of utility or value, or that it treats persons as mere containers for value. When faced with such arguments, consequentialists often claim that this mistakes the consequentialist conception of value–that utility is not some thing over and above persons, but rather that we care about utility because we care about persons. I discuss this further in "On the Cross of Mere Utility."

the price and deny that there is any person-centered ultimate rationale for promoting the good. But doing so leaves them with some very difficult questions to answer: Why should we (we persons) care about the quantity of some (literally) impersonal value? What account can we give for the value of states of affairs that do not presuppose the value of persons or other sentient beings? And what rationale can we give for the claim that certain things–other than persons and other sentient beings–simply have value that is in no way derivative of the value of persons or their capacity to bestow, create, or project value into the world? I think that it is fairly difficult to see how such questions could be answered without an appeal to brute intuition. But of course if brute intuition is all that supports the impersonal theory of value, then it is difficult to see why that intuition is to be preferred over the widespread intuitions that persons are a fundamental source of moral value and that one of the main roles of morality is to guide us in responding to that value.

As we noted at the outset, moral philosophers since the ancient Greeks have assumed that there must be some connection between a moral theory and the beings whom it is to govern. The ideas that a moral theory should express concern for persons in a way that responds appropriately to what they are like, and that the requirements that it gives must be able to guide the beings to whom they are given, are not simply ad hoc intuitions brought in by those who have a particular moral theoretic axe to grind. Instead, they are part of the common heritage of Western moral philosophy from its very beginnings. Because they are so deeply embedded in the history of moral philosophy, it is arguable that they form part of the very concept of morality that has emerged from our common intellectual tradition. Indeed, a theory about anything must display a certain continuity with our commonsense understandings of the subject matter if it is to count as a theory about the same thing that we set out to theorize about. If a moral theory departs too much from our common understanding of what morality is, then it is unclear that it really is a theory about the thing we have in mind when we use the term "morality."[44]

---

[44]Notice that the major competing method of moral inquiry actually requires much more in terms

It is always possible, of course, that our traditional understanding of morality should be revised. But notice that the idea that morality is fundamentally (though perhaps not exclusively) concerned with persons is not only part of our common understanding of what morality is, why we care about it, and how it addresses the human condition, but it is also an eminently reasonable assumption that cannot be easily impugned as mere prejudice or superstition in the way that some of our less savory traditional ideas can be. It certainly seems reasonable to place the onus on those who wish to reject this understanding of the nature of morality. That burden cannot be carried simply by noting that the understanding is at odds with one's favorite moral theory.

In the end, those who use the nature to morality methodology may not have an airtight argument for it. (But in moral philosophy, who does?) However, they do get the luxury of saying that when they do moral theory, they are doing the same sort of thing that moral philosophers have done as far back as Aristotle. And they can take some comfort in the fact that those who reject the nature to morality methodology are doing something quite different. Even if the exploration of human nature and the nature of persons does not, in the end, settle completely the debates between competing moral theories, the very project would have what I would regard as the salutary effect of bringing moral philosophy back to its roots. It would constitute an explicit recognition of something that our intellectual ancestors took for granted: that a moral philosophy that speaks to the human condition cannot be completely separated from inquiries into human nature or the nature of persons.[45]

---

of our taking our traditions for granted. The method of reflective equilibrium requires that the content of a theory not depart too radically from the content our ordinary moral thinking. The nature to morality methodology, on the other hand, requires only continuity with a conception of the nature and point of morality, i.e.,the conception of morality as being for and (at least mainly) about persons.

[45]I am extremely grateful to anonymous reviewers and an editor for Canadian Journal of Philosophy for extremely generous and helpful suggestions. I am grateful to Bob Stecker for

NOTES

---

several helpful comments and Shelly Kagan for some early words of encouragement on this project. Much of my thinking about this topic took shape while I was teaching Scheffler and Kagan in courses that I taught at Simon Fraser University in the fall of 1997 and at Central Michigan University in the fall of 2000. I thank both of these truly wonderful groups of students, especially Tim Christie and Lou Blouin, for conversations that helped me to work out some of the ideas contained here.