

## Two Sides of the Meta-Ethical Mountain?

Peter Railton

[written for a planned volume, *Does Anything Matter? Parfit on Objectivity*, ed. by Peter Singer]

### Introduction

Derek Parfit's *On What Matters*<sup>1</sup> is a book that matters. It is a contribution to systematic ethics of such sweep and cumulative argumentative force that it will have a central role in debates in moral philosophy for years to come. There is a great deal in *On What Matters* that I find compelling. In an age marked by rather stark oppositions between “Kantian” and “Humean” approaches to ethics, or “Deontologist” and “Consequentialist” normative theories, *On What Matters* achieves a remarkable synthesis. For it suggests that these approaches, properly understood—and therefore liberated to some extent from a number of their contemporary advocates—might be seen as climbing different sides of the same mountain, to find themselves united at the summit.

In meta-ethics, however, Parfit sees things differently. Naturalists never get beyond the foothills of normativity. When they try to scale the heights, the bottomless crevice of Nihilism awaits them, not re-union with the Parfitian Non-Naturalist: “Normativity is either an illusion, or involves irreducibly normative facts” (267). Were Naturalism to be our only option, Parfit writes, much of his philosophical life, as well as much of the lives of such great modern moral philosophers as Sidgwick and Ross, would have been “wasted” (303-304). Worse, “we would have learnt that nothing matters” (367).

This sad thought is made all the more poignant by the fact that Parfit sees the situation as asymmetric. Even if they prove wrong on matters meta-ethical—as it seems we must hope they are—the great Naturalists of the past such as Hobbes, Hume, Bentham, and Mill will not have wasted *their* philosophical lives: they will have taught us something about the natural *reason-giving* features of the world. That is, the Naturalists might have described everything that is of genuine *normative importance*—pleasures and pains, well-being, Mill's “permanent interests of man as a progressive being”, and so on—even though the nature of normativity itself altogether eluded them. The non-natural *normative properties and facts* Parfit insists upon, he explains, themselves have no such importance. They serve to formulate reasons-relations, but do not give “any *further* reason” for acting in their own right (279-280).

---

<sup>1</sup> In 2 Volumes. Oxford: Oxford University Press, 2011. Parenthetic page references in the text, unless otherwise indicated, are to Volume Two.

## The “decisive battlefield”

Thus, we might be aware of pleasure and pain, weal and woe, hope and despair, helping and harming, truth and falsehood—everything of that is of normative importance—yet absent recognition of independent reasons relations, “that importance would be unknown to us—as it is unknown, for example, to some active, intelligent cat” (288). Were I to say, ‘You would enjoy this book’ or ‘Your wine is poisoned’, this would “tell you facts that would give you reasons for acting”, which you might take to heart and be glad you did, but this could not constitute genuine *advice*—for that we need the irreducibly normative concept of *reasons for acting* (281).

Of course, many Naturalists would accept the existence—and usefulness—of irreducibly normative concepts, precisely for such purposes as reflecting individually or together on what matters, and deliberating, deciding, and advising about what to think or feel or do. They also accept the existence of substantive truths involving these concepts. Parfit calls such views “Soft” Non-Analytic Naturalism, and Parfit himself has sometimes seen the initial plausibility in such a position, despite his long years of adamant resistance to Naturalism (349). Moreover he allows that Naturalists of various kinds might be able to give a plausible account of such forms of normativity as the normativity of correctness (including truth or falsity for belief or correctness for assertion, 266), rules, customs, laws, meaning, well-being, and even beauty, rationality (in the procedural sense), and moral rightness (in the attitude- or practice-based sense)—indeed, as these notions are often understood (267-269).

But explaining these is not the “decisive battlefield” in meta-ethics (269). Nor is the battle over what we ordinarily have in mind when making normative judgments. Rather, it concerns what we *should* have in mind, at least, when thinking at the most fundamental level of normative reflection (272). And that is the idea of unqualified reasons—reasons *full stop*. It is ‘reasons’ in this sense that enable us intelligibly to ask, of any of these other normative categories, *Do we, in the end, have sufficient reason to take them seriously—to guide our lives by these rules, procedures, attitudes, or practices, or to hold ourselves to these standards of correctness?* This is the “*reason-involving*” or “*reason-implying*” sense of normativity, the sense in which it implies the existence of “some reason or apparent reason” (268). And this is where the buck stops when we are asking what really matters—or seeking decisive guidance as to what we ought to think and do.

Thus, a Naturalist might be able to give an account of morality in something like the sense in which Hume understood it (310)—the sense in which it is embodied in a constellation of human attitudes, traits of character, norms, practices, and concerns as grounded in relationships of mutual cooperation and shared discussion and restraint, which promote the “general interest” and earn our reflective approval and motivated commitment by making our lives go better. But accepting that something is moral in this sense does not logically imply that I have sufficient reason to take its values or injunctions as action-guiding. I might decide that I

should do so, even that it would be a terrible mistake not to, but then in giving this answer I am to be understood as answering a normative question neither posed nor precluded by the original standards themselves. Fortunately, Parfit believes that there are *objective, external* reasons, independent of what we happen to care about or find motivating, and that these reasons can speak on behalf of morality. And because they are reasons *full stop*, I cannot in a similar way ask a meaningful normative question about whether I ought to heed them. For to ask such a question is simply to ask whether I have sufficient reason for so doing.

Now it might seem that we *could* ask such a question, precisely because these reasons are external. We might see that an act is supported by external reasons, yet find that these reasons simply do not speak to us and our deepest concerns. But regarding what actually or reflectively moves us, or is endorsed by us, Parfit writes:

There is, I believe, no normativity here. An irresistible impulse is not a normative reason. Nor is an impulse made rational by its ability to survive reflection on the facts. Even after carefully considering the facts, we might find ourselves irresistibly impelled to act in crazy ways. [291]

After all, someone might find that bloody revenge is what speaks to him—even after full reflection, and even at the acknowledged cost of any hope of future happiness for all concerned. If so, and if he exacts vengeance as a result, he is acting with no reason and making a “*terrible mistake*” (437, 292). “There is something else, and something better, for normativity to be” (285). Indeed, even the hard-core “Internalist” who disagrees is in fact making a contrary “external” or objective claim—a claim about what reasons there are, *full stop*. Parfit concludes, “To avoid confusion, we should use the phrase ‘a reason’ only in its external, irreducibly normative sense” (290).

### **Concepts and properties: the Soft Naturalist’s Dilemma**

Parfit argues at length that, at the core of the dispute between the Non-Naturalism about reasons he favors and Non-Analytic, Soft Naturalism—hereinafter what I will mean by ‘Naturalism’ when used without qualification—lies in a seemingly rather subtle question. Namely, whether statements like:

(A) What we have most reason to do is to maximize happiness

admit a reading in which they can coherently be understood as synthetic property identities, in which two concepts, one normative (<most reason to do>) and one natural (<maximizing happiness>), pick out one and the same natural property, *maximizing happiness*, without any

intermediation by a non-natural property.<sup>2</sup> Such a reading would permit the Naturalist to accept that normative concepts cannot be reduced to natural concepts, while at the same time having no need to enlarge her metaphysics beyond natural properties. The property *maximizing happiness* would do double duty, descriptive/explanatory and normative. In consequence, the Naturalist's ontology would not need to admit irreducibly non-natural properties. On such a reading, a Naturalist who adhered to (A) could be saying that an act's rightness "consists in" its maximization of happiness, or that, to maximize happiness "is what it is for" an act to be right. (A), if true, would be a synthetic, substantive truth, discovered *a posteriori*.

An analogy would be with a "Soft" Materialism according which, for example, the property of *having C-fibers firing* can do double duty, figuring in physiological explanations of neurological and behavioral phenomena, on the one hand, and as the referent for the mental concept <experiencing pain> and truth-maker for pain ascriptions, on the other.<sup>3</sup> This would allow the Materialist to recognize the distinctive inferential and conceptual roles of mental ascriptions—situating talk of 'pain' within the commonsensical mentalistic conceptual framework. It would also account for the *a posteriori* nature of the discovery that pain consists in C-fiber firing, while at the same time positing no irreducibly mental properties or substances.

Another analogy would be with a "Soft" Physicalism according to which the microphysical property of a system *having a certain mean molecular kinetic energy* can do double duty, figuring in thermodynamical explanations of the system's behavior, on the one hand, and as the referent for (what physicists call) the phenomenological concept <having a certain heat> and the truth-maker for heat-attributions, on the other.<sup>4</sup> This would allow the Soft Physicalist to recognize the distinctive conceptual and inferential roles of heat attributions—situating talk of 'heat' within a macroscopic, observational conceptual framework. It would also account for the *a posteriori* nature of the discovery that phenomenological heat consists in microscopic kinetic energy, while at the same time positing no irreducibly non-physical properties or substances.

Parfit believes that these Soft views all are incoherent. They face a kind of dilemma, in which the only form in which claims like (A), or

(B) To be in pain is to have one's C-fibers firing,

(C) To have a certain degree of heat is to have a certain mean molecular kinetic energy,

could be true and informative—as they are intended to be—is a reading according to which they involve a relation between *two* distinct properties. That is:

---

<sup>2</sup> I will use the convention of italicizing *properties* and placing <concepts> in corner braces. (A) is unrealistically simple, but it will serve for illustrative purposes.

<sup>3</sup> This example is unrealistically oversimple, but it will serve for illustrative purposes.

<sup>4</sup> This, too, is unrealistically oversimple.

(A+) Whatever act has the normative property of being what we have most reason to do has the different, natural property of maximizing happiness.

(B+) To have the mental property of being in pain is to have a different, physical property of having C-fibers firing.

(C+) To have the phenomenological property of possessing a certain degree of heat is to have a different, microscopic property of possessing a certain mean molecular kinetic energy.

Here is how the argument goes.

On one horn of the dilemma, the Soft Naturalist (or Materialist, or Physicalist) sees herself as advancing a substantive, informative, and positive claim in asserting (A) (or (B) or (C)). Let us take (A), and suppose this claim to be right. What sort of claim is (A) and what would it take for such a claim to be substantive, informative, and positive?

Could (A) be a purely non-normative claim? No, since putting this claim forward has implications about a whole range of questions about what we have most reason to do, and that is what normativity is all about. Someone dissenting from (A) on the ground that following the Golden Rule is what we have most reason to do, would have a normative disagreement with (A). So (A) has non-trivial normative content.

When is a normative claim substantive? Parfit offers us this definition:

[def. 1] Some normative claim is *substantive* when this claim both (a) states that something has a normative property, and (b) is *significant*, by being a claim with which we might disagree, or which might be informative, by telling us something we didn't already know. [275]

And when would such a claim significant or informative? Parfit gives these quasi-definitions or explications:

[def. 2] Any such information must be statable ... as the claim that such acts would have one or more other, different properties. [344]

[def. 3] As I use the concept of a *property*, any information about such acts could be stated as the claim that these acts would have some property. [348]

And finally, when would it be positive?

[def. 4] ... normative claims are *positive* when they state or imply that, when something has certain natural properties, this thing has some other, different normative property. [343]

But now it is clear that, given these definitions, (A) can be substantive, significant, and positive only when it is interpreted as:

(A+) Whatever act has the normative property of being what we have most reason to do has the different, natural property of maximizing happiness.

And yet (A+) is plainly incompatible with the Soft Naturalist's claim that only one irreducible property, the natural property *maximizing happiness*, need be in play in (A).

Thus, the effort to make the Soft Naturalist's position something worth asserting—a substantive, informative, positive claim—makes it false. A similar argumentative strategy works, *mutatis mutandis*, to show that (B) and (C), if they are to be substantive, informative, positive claims, must be interpreted as (B+) and (C+).

On the other horn of the dilemma, the Soft Naturalist insists upon something like a “referential” reading of (A), so that it reaches directly to the underlying natural property, with no recourse to an intermediating normative property. Let us grant this reading. But then stating (A) in effect equates the reference of <having most reason to do> with the reference of <maximizing happiness>. Yet this is to say nothing more than:

(A\*) Whatever act has the property of maximizing happiness has the property of maximizing happiness.

And that is trivially true. *Mutatis mutandis*, (B) and (C) would say nothing more than:

(B\*) To have the property of having C-fibers firing is to have the property of having C-fibers firing.

(C\*) To have the property of having a certain mean molecular kinetic energy is to have the property of having a certain mean molecular kinetic energy.

So, on the “referential” reading, (A)-(C) would all come out true, but as thinly disguised tautologies, not substantive, informative, positive philosophical positions.

Thus, the effort to make Soft Naturalism (or any of its Soft mates) true makes it trivial, not substantive. Since we have already seen that the effort to make it non-trivial makes it false, the dilemma appears to be complete.

If Soft Naturalism is thus incoherent, and normative concepts must be seen as referring to normative properties—even though normative concepts may *also* refer to natural features *via* referring implicitly to normative properties (331)—then to adhere to a Naturalistic metaphysics that *excludes* unreduced normative properties forces the would-be Naturalist to be a “Hard” Naturalist, holding an Error Theory about normative claims. Thus Naturalism is “close to” Nihilism (368)—where normativity is seen as an illusion and “nothing would really

matter” (cf. 367). Similarly, Hard Materialism would lead to an Error Theory about the mental (e.g., in reality, there are no pain experiences, just physical states) and Hard Physicalism to an Error Theory about the phenomenological (e.g., in reality, nothing is hot or cold, or red or blue, there are just microphysical differences in energy level, or colorless light-reflectance properties).

These are surprisingly powerful results. If the Naturalist were attempting to scale the same meta-ethical mountain as Non-Naturalists like Sidgwick, Ross, and Parfit, she would just have tumbled off a ledge into normative oblivion.

## Types of Naturalism

There are many steps in the arguments rehearsed above against Soft doctrines.<sup>5</sup> How are we to assess them?

First, how did we get to the conclusion that (A+)-(C+) were the mandatory interpretations of (A)-(C) when these are substantive, significant, and positive? Via the definitions or quasi-definitions [def. 1]-[def. 4], repeated below:

[def. 1] Some normative claim is *substantive* when this claim both (a) states that something has a normative property, and (b) is *significant*, by being a claim with which we might disagree, or which might be informative, by telling us something we didn't already know. [275]

[def. 2] Any such information must be storable, however, as the claim that such acts would have one or more other, different properties. [344]

[def. 3] As I use the concept of a *property*, any information about such acts could be stated as the claim that these acts would have some property. This property might be linguistic. [348]

[def. 4] ... normative claims are *positive* when they state or imply that, when something has certain natural properties, this thing has some other, different normative property. [343]

A philosopher should perhaps be allowed whatever definitions or quasi-definitions he wants; the chief interest lies in what they enable him to establish, so long as they don't beg relevant questions. But [def. 1]-[def. 4] have the interesting feature that they all introduce 'property' or 'normative property' or 'different property' into the definiendum or quasi-definiendum. Unsurprisingly, then, they together have the result that (A)-(C) won't be substantive,

---

<sup>5</sup> We will come to an argument peculiar to the Soft doctrine in normative case, the "Normativity Argument", below.

informative, and positive *in Parfit's senses of the terms* unless the predications they contain are interpreted in terms of the attribution of different properties, that is, as (A+)-(C+).

To block this, we need only find coherent, alternate ways of explicating how (A)-(C) could be *in some other, philosophically interesting sense* substantive, informative, and positive, even if Soft Naturalism, Materialism, or Physicalism were true.

Why bother? Perhaps Soft Naturalism, Materialism, and Physicalism are radically implausible doctrines, the brain-children of metaphysically obsessive-compulsive philosophers who dogmatically insist at the outset that:

- (G) All irreducible properties or facts are natural.
- (H) All irreducible properties or facts are material.
- (I) All irreducible properties or facts are physical.

Parfit writes of “Normative Naturalists” that:

... most are also *Metaphysical* Naturalists, who believe that all properties and facts are natural, so that there could not be any irreducibly normative facts. [267]

But it seems to me that Hobbes, Hume, Mill, and a great many historical and contemporary naturalists, myself included, are motivated primarily by *methodological* as opposed to *a priori metaphysical* considerations. If they resist non-natural entities, properties, or forces, it is because they see them as problematic in the effort to conduct fruitful, empirically-disciplined inquiry that leads to the development of genuinely explanatory theories.

Methodological Naturalism is the view that, in inquiry, we should try to proceed by following the ways of observing, hypothesis-forming, testing, formal modeling, axiomatizing, and explaining that have been so successful in the natural, biological, social, and mathematical sciences (Railton, 1997). Hobbes and Hume, for example, saw themselves as applying emerging methods of scientific inquiry to develop an empirically-based “science of man”, not trying to duplicate the dogmatic metaphysics of their predecessors, only with a Naturalistic flavor.

Methodological Naturalists do not begin with bald metaphysical assertions like (G)-(I), since their approach to metaphysical questions has to be: “Let’s see”. That is, “Let’s see how far we can go with the hypothesis that all facts or properties are of a kind that figure in the forms of inquiry and explanation found in the sciences—or can be understood in terms of conceptually-clear constructions from such facts and properties”. It could turn out that there is a great deal that science cannot account for, even granting that every mode of inquiry is allowed to let explanations stop somewhere. Prior to Darwin, for example, religious and irreligious Methodological Naturalists worried how to explain the extensive evidence of “design” and “pre-established harmony” to be found in nature. Were there no evidence for

evolution or other possible mechanism of natural selection, then a methodological preference for non-coincidental explanation would leave us today in the situation described by Hume, “Those, who delight in the discovery and contemplation of *final causes*, have here ample subject to employ their wonder and admiration” (Hume, 1888, IV.ii.44).

More recently, methodological considerations led to a metaphysical shift within the Naturalist camp. Many naturalistically-inclined philosophers who had been drawn to Central-State Materialism shifted to some form of Functionalism once the development of programmable computers and evolutionary psychology convinced them of the greater generality and explanatory power of Functionalism. They were Methodological rather than Metaphysical in their approach to the question of Materialism, so that (B) was not an *a priori* commitment or starting point, but a substantive *a posteriori* identity claim—one that now seemed, for good methodological reasons, to be false.

What has tended to drive Methodological Naturalism is an effort to be modest about claiming any special philosophical insight into the nature of reality—whether it be the reality of commonsense, or of the nature of mind and meaning, or of morality, or of fundamental metaphysics. From such this perspective, a defeasible preference for what David Lewis (1983) called “sparse”—as opposed to “abundant”—views of properties is not arbitrary. The introduction of *any* irreducible properties, entities, or facts raises fresh questions about intelligibility, epistemic and semantic access, coherence with existing theory, explanatory gain, and appropriate forms of discipline in attribution. These worries become acute if they involve synthetic claims stretching to putative metaphysical features outside the causal order. Parfit frankly allows that “when I say that we have some reason, ... what I mean cannot helpfully be explained in other terms” (272), and that, in contrast to prosaic natural language, it is more difficult to explain “how we come to understand [normative] words and the concepts they express, and how we can recognize any irreducibly normative truths” (293).

He correctly notes that this is not decisive, and later in the book has a number of interesting things to say about evolution and possible non-causal access to normative facts (507ff). But it is hardly arbitrary to think that these explanatory problems might speak in favor of a sparser account of the properties and facts involved in normative discourse. If we already countenance both naturalistic *reason-making* properties and facts and the irreducibly normative *concepts* needed to discuss and deliberate upon them, one might well ask, is this sufficient for understanding normativity? Moreover, supervenience guarantees that natural facts will suffice to draw whatever genuine normative distinctions might exist among acts or outcomes. For there can be no genuine normative differences without genuine natural differences.

## Concepts and properties, second pass

Let us now return to the problem of showing whether (A)-(C) could be given a coherent interpretation that preserves the “Soft” idea that normative, mental, and phenomenological concepts might be irreducible, even if normative, mental, and phenomenological properties, states, or facts are not. Are there credible alternatives to (A+)-(C+)?

Historically, few topics have led to the spilling of more philosophical ink than whether properties are needed to explain predication, and if so, when and in what sense? Nominalists, Conceptualists, Tropists, Fictionalists, Minimalists, and other deflationary theorists hold that statements like (A)-(C) can be understood without recourse to properties as irreducible, abstract elements of reality. Some treat properties in a nominal or linguistic sense (e.g., as “the shadow of a predicate”). Others analyze properties in terms of classes or sets of individuals, actual or possible, worldly or trans-worldly or centered-worldly, and see predication in terms of membership in such classes or sets. And so on.<sup>6</sup> It would be surprising if all of these views were incoherent, so that interpretation of (A)-(C) in terms of irreducible abstract properties were mandatory.

Moreover many philosophers, some moved by methodological concerns about semantic or epistemic access, have pursued *selective* forms of Nominalism, Fictionalism, or other forms of deflation. For instance, some recently have sought to show that metaphysical commitment to abstract mathematical entities and properties, lying outside of the causal order, can be avoided without giving up elements of the formal or conceptual apparatus of mathematics needed in science (Field, 1993; Yablo, 2001). Others have pursued Fictionalism about the normative as an alternative to Expressivism (Kalderon, 2005).

Parfit might argue that such forms of Nominalism, Fictionalism, and the like are not plausible, or not uncontroversial. But the question is, are they simply incoherent? As Parfit points out in his discussion of “speech act” arguments that purport to derive *ought* from *is*, the mere intelligibility of Act Consequentialism as a normative view, regardless of its truth or falsity, shows that these inferences cannot go through without adding some substantive normative claim (310-314). Similarly, we might say that the mere coherence of even one of Nominalism, Fictionalism, Tropism, etc., shows that Parfit’s attempt to show that (A)-(C) must be read as (A+)-(C+) requires some further assumptions involving properties. Indeed, [def. 1]-[def. 4] could be seen as reflecting just such assumptions. On a natural reading, [def. 3] corresponds to an “abundant” view of properties. Moreover, it would appear to take sides on the

---

<sup>6</sup> Similar remarks apply to deflationary views about *facts*.

controversial question whether all information—including, say, the essentially indexical and practically vital information that an act is *one's own*—is “statable” as a property.

Defenders of Nominalism, etc., about non-natural, immaterial, or non-physical properties—whether as part of a general deflationism about properties or as a “sparse” view of properties that accepts natural kinds as *bona fide* properties, with an explanatory role to play—give philosophical reasons for preferring an account of statements like (A)-(C) that do not invoke irreducible non-natural, immaterial, or non-physical properties. Whether to accept any such interpretation seems to me more a matter for philosophical cost/benefit calculation, as Lewis would say (Lewis, 1986), than a situation in which one side refutes the other.

To get a feel for costs and benefits, consider how Parfit's preferred account analyzes statements akin to (A)-(C). On his account, macroscopic ‘heat’ means:

[i] the property, whichever it is, that has the *different*, second-order property of *being the property that can have certain effects*, such as those of melting solids, turning liquids into gases, etc. [330]

Therefore,

When scientists discovered that

[ii] heat is molecular kinetic energy,

what they discovered was that

[iii] molecular kinetic energy is the property that has this *different*, second-order property. [331]

But it seems odd to ascribe causal effects like melting solids to an abstract second-order property that is, moreover, distinct from the first-order physical property of possessing a certain molecular kinetic energy. Is such metaphysical ascent or reification really needed to understand (C)?

Here is an alternative, de-reifying approach. Think of *causal efficacy* as residing in actual or potential spatio-temporal individuals or systems. *Causal explanations* then can be given in terms of properties that carve the world into classes of individuals or systems along the lines of similarities and differences in such causal efficacy—“natural kinds”. Since such properties are not something over and above these classes of individuals or systems with similar causal powers, invoking them explanatorily does not suggest that they make any *independent* contribution to causal effects. They therefore raise no naturalistic objections.

A phenomenological concept like <possessing a certain degree of heat> has a distinctive place in our conceptual framework, and we can explicate this in terms of a complex role or “job description” to be satisfied by whatever would satisfy this concept—involving a capacity to melt solids, turn liquids into gases, etc., but also to do work, to produce certain characteristic sensations (“hot” or “cold”) in observers under standard conditions, to correspond to temperature differences and changes (“heat flow”), etc. Such a “job description” need not be a *conceptual analysis* properly so called, since it involves a great deal of substantive empirical content and since any of its particular elements—detailing conceptual interrelations, causal or constitutive connections, paradigm cases, etc.—might ~~turn~~ be revised or abandoned without leading us to abandon the phenomenological concept <heat> altogether. Some early modern theorists would have included substantiality in the job description, but this was given up when the kinetic theory emerged. Similarly, phenomena like sensations of hot and cold and heat transfer were discovered to be influenced by *density* and *conductivity* as well as *heat*. But neither of these theoretical developments led to dropping the concept <heat>. It is this complex job description that philosophers and historians of science tease out when they try to reconstruct the concept of <heat> as it has functioned and evolved in commonsense and scientific discourse and practice.

We now can say that (C) can be interpreted as saying, in effect,

(C•) Collections of molecules possessing a certain mean kinetic energy uniquely fulfill the job description associated with satisfying our phenomenological concept <possessing a certain degree of heat>.<sup>7</sup>

(C•) is plainly a substantive, informative, positive claim. Certainly it was non-trivial news as it emerged over the course of experimentation and the development of thermodynamics. It is quite distinct from the claim:

(C\*) To have the property of having a certain mean molecular kinetic energy is to have the property of having a certain mean molecular kinetic energy.

(C•) thus is enough to enable a Soft Physicalist to explain the substantive, informative, positive character of (C), without positing irreducible phenomenological properties, and while allowing that phenomenological concepts cannot be reduced to physical concepts. Parallels in the case of mind, in which mental concepts are associated with psychic and “folk explanatory” roles,

---

<sup>7</sup> In a serious treatment, (C•) would be strengthened to a claim of necessity, and would be a matter of molecular kinetic energy *best* or *sufficiently* fulfilling this role. Once these qualifications are spelled out, the option becomes available of using such roles or job descriptions to provide a kind of “non-obvious” definition or analysis (Lewis, 1970; Railton, 1993). For reasons suggested above, as well as independent reservations about what “Ramsification” in general can establish, it seems to me preferable not to see them this way.

which then are realized by neurophysiological or functional systems, are well known. The Soft Materialist's interpretation of (B) might be:

(B•) C-fiber firings in the central nervous system uniquely fulfill the job description associated with satisfying the mentalistic concept <pain>.<sup>8</sup>

## Normative concepts and normative roles

Let's return to the normative case. Can we use the approach just discussed for (C) to explain the substantive, informative, positive character of (A):

(A) What we have most reason to do is to maximize happiness.

As we saw, on Parfit's view, there is not a lot to say about the concept of a reason *full stop*. On the other hand, he certainly does think that <most reason to do> has a distinctive, all-important role in our conceptual scheme—it expresses that which ultimately matters, or appears to do so, and “stops the buck” in deliberating and deciding what we ought to do. Moreover, he also thinks there are paradigm cases of what is relevant to <most reason to do>, or what it includes or excludes, which recur throughout the book: future pain and well-being count, agony or vengeance that serves no further purpose do not, and so on.

If we did not have at least this rather “thin” idea of the role of <most reason to do> in thought and deliberation, we would be hard put to learn it, teach it, or distinguish it inferentially from other primitive concepts that can be applied to acts, like <actual> or <possible>. Why would attaching <most reason to do> to an act have any relevance in deciding what to do different from attaching <possible>, if both were simply inexplicable simples with no articulable roles? How would you know I had mastered the concept <most reason to do> if there were no paradigm cases, and the only examples I could give after careful reflection were *skipping rope* and *bragging*?

Parfit's view of the role of <most reason to do> may be “thin”, but for those who think that some form of internalism is part of this concept—reasons must be the sort of thing *for which* one can act, that could engage positive affective interest and motivation or gain considered endorsement without further incentive, and so on—the associated job description would be considerably “thicker” and richer.

Given some such relatively “thin” or “thick” role, how might (A) be interpreted by the Soft Naturalist as substantive, informative, and positive?

---

<sup>8</sup> See the qualifications for (C•), above.

(A•) Acts that maximize happiness uniquely fulfill the job description associated with satisfying the normative concept <most reason for action>.<sup>9</sup>

(A•) is plainly more informative than:

(A\*) Whatever act has the property of maximizing happiness has the property of maximizing happiness.

That (A•) is true would come as news to a great many normative inquirers, past and present, while (A\*) would not. Moreover, (A•) is normatively substantive and positive, since one would learn something positive about what one ought to do in learning it, e.g., that *maximizing happiness* always stops the buck in deciding what one ought to do. Someone who thought that only the Golden Rule had this status would urge instead:

(J) Acts that conform to the Golden Rule uniquely fulfill the job description associated with satisfying the normative concept <most reason for action>.

And a Nihilist or Error Theorist would urge:

(K) No acts fulfill the job description associated with satisfying the concept <most reason to do>.<sup>10</sup>

(A•) takes substantive issue with (J) or (K). Moreover, if true, it would have the status Parfit thinks normative truths must have, for, fully spelled out, it would formulate a necessary but non-analytic claim, as would (J) or (K). Since the role specified by <most reason to do> will essentially involve other normative concepts—e.g., <ought to do> in the “thin” role, also <ideal conditions> in the “thicker” role—this role cannot be reduced to a naturalistically-characterizable role, and so would not afford a naturalistic reduction of the concept.<sup>11</sup>

As far as I can see, (A•) is as coherent as (B•) or (C•).

Parfit will of course prefer a translation of (A•)-(C•) into a language of properties and higher-order properties, e.g., “the property *being such as to fulfill the role ...*, is had by whatever has the property ...”, but as far as I can see this is a philosophical choice, not mandated by the need to distinguish (A) from (A\*), since (A•) does that. Moreover, (A•) does this without imputing causal features to second-order abstracta and it is metaphysically less profligate, since conceptual roles come along with concepts, but for “sparse” thinkers, (A•), like (B•) and (C•),

---

<sup>9</sup> See the qualifications for (C•), above.

<sup>10</sup> Again, see the qualifications on (C•), above. Since Mackie suggests that perhaps the existence of God would make “objective prescriptivity” possible (1977), his Error Theory would take the form of asserting that “No acts fulfill the role specified by the concept <most reason to do> in any Godless world”, along with “Our world is Godless”.

<sup>11</sup> For an alternative view, Jackson (1998).

can be accepted without invoking basic properties that go beyond the natural kinds delineated by fundamental causal/explanatory similarities.

Indeed, I wonder whether there might be less difference than at first seems to be the case between Parfit's view and the Soft theories presented here. Parfit's definitions and quasi-definitions have the effect of generating a new property wherever there is new information. Normally, it is not *properties* but *properties under descriptions* that carry information, but since the property might be linguistic [def. 3], this distinction is blurred. Now, even strict Nominalists can tolerate *linguistic* properties, so long as they do not involve adding anything to their *ontologies* beyond individuals—including linguistic entities—and classes thereof. And Parfit says that non-natural properties exist in a *non-ontological* sense (482). Moreover, as we saw at the outset, Parfit believes that normative properties and the normative facts involving them are, in themselves, of no *normative importance* (279-280), and so do not figure in any fundamental metaphysical explanation of what really matters.

Soft Naturalists can accept *non-natural* properties in a nominal or linguistic sense, so long as they not asked to do fundamental explanatory work. And they allow us to talk meaningfully and truthfully in terms of normative concepts—making reasons claims in which *normative predicates* figure. And true claims of this kind can be called by the Naturalist in one familiar sense of this term, *normative facts*. Of course, the Naturalist of this kind will say that such facts make no contribution in their own right to the fundamental explanation of what has normative importance—but Parfit isn't saying that they do. Parfit moreover does not imagine that we have some special, quasi-perceptual mode of access to normative properties or facts—ordinary perceiving, thinking, and acting in the world suffice.

Reflecting on all this, I'm not sure how strenuously a Soft Naturalist should object, if at all, to “non-ontological” non-natural properties or to the normatively unimportant non-natural facts attributing them. Soft Naturalists surely object to Platonistic Non-Naturalism, complete with a distinctive, epistemically-privileged faculty of *a priori* intuition—but these are *ontological* non-natural properties, and that way of knowing synthetic truths is *sui generis*.

Of course, there is much more to be said. This is only an attempt to lob the ball back into Parfit's court in the wake of his arguments against Soft Naturalism, and to invite a return volley.

### **“Undeniably in different categories”—the Normativity Objection**

But Parfit raises another objection that suggests the Soft Naturalists' lob goes wildly out of bounds. And in this case, it pertains specifically to *normative* properties and facts, not properties and facts in general. Parfit allows that Naturalists *could* be right about many

normative matters, as we noticed earlier. Moreover, as we noted, he allows that the Naturalists' particular proposed candidates for the natural properties picked out by normative terms might be acceptable as accounts of what normatively matters—natural properties can *make* an act be wrong, or what we have most reason to do. Parfit's "Normativity Objection" (324-327) is that none of these natural facts or features could possibly be what *being wrong* or *being a reason* could consist in. They are "undeniably in different categories"—"Rivers could not be sonnets, experiences could not be stones", and "Natural facts could not be normative in the reason-implying sense" (324-325).

Given the differences between the concepts involved, he believes, it simply is not "left open" that rivers could turn out to be sonnets, or stones experiences, or natural facts normative facts. So we cannot even entertain such identities coherently—the attempt to accept them results in nonsense. This in itself would suffice to rule out Soft Naturalism, even setting aside all that has gone before.

But there is a crucial difference between the case of sonnets and rivers or experience and rocks, on the one hand, and the case of natural and normative facts, on the other. Rivers do not supervene upon sonnets, nor experience upon rocks. All the sonnets in the world do not determine which rivers there are, nor do all the rocks determine which experiences. But all the natural facts of the world, taken together, *do* suffice to determine (metaphysically, not analytically) all the normative facts.<sup>12</sup> This is an important feature of normative facts. Some would say that it is part of our *a priori* understanding of value or the normative that it is supervenient upon the non-normative in this way.

And that makes sense. After all, as Parfit agrees, the things of normative importance include plainly natural things like pain, happiness, accurate belief, and so on. Fix these things, and you have fixed what the reasons-making facts are. Not so with sonnets and rivers, or stones and experiences. We can imagine a God-like being arranging all the features of all the stones in the world, then wondering whether to add to this world any consciousness—perhaps, to appreciate Her handiwork. But we cannot imagine such being fixing all the features of an act of inflicting prolonged agony upon another for one's own amusement, then having a choice to make about whether to add any normative importance to whether this occurs or continues, or whether it is to be a matter of value or disvalue. Here, I think, we do encounter a genuine inconceivability.

Or think of what it would be like trying to act well if all that one could do by way of causal efficacy—carefully attending to those in pain, respectfully helping them escape their torment and go on to lead fulfilling lives—would still have to await some further something (what?) in order to be worth doing. Someone else might act identically to you, and he would

---

<sup>12</sup> As far as I can see, Parfit discusses supervenience only in an endnote (300n), where he comments simply that normative supervenience needs separate treatment for the supervenience of the mental.

merit moral praise while you might be wasting everyone's time or doing something heinous. Normative judgments of what we ought to do can do their job of guiding action because what we *can* affect is the very stuff that makes for goodness or badness.

Whatever value is, and however it might be different from fact, this profound connection between the natural and the normative is central to it. That makes the fact/value relationship more intimate than sonnets to rivers, or even of complex neurological states and consciousness (the supervenience of the mental not being in the same way part of our mental concepts). More strongly, as Moore thought, value is "consequential" in relation to the natural—there is an *asymmetric dependence*. This suggests there is something metaphysically *explanatory* going on, with the natural having primacy.

Mere supervenience commits us to the existence of some natural property—perhaps complex and disjunctive and uninteresting *qua* natural—that holds for all cases in which a given evaluative concept applies. But the distinctive asymmetric dependence of the normative upon the natural suggests something stronger. For example, grasping the conceptual role associated with the concept <reason for action> involves some notion of "*ought* implies *can*", and someone would have a defective understanding of the concept <intrinsically valuable> if he found it mystifying why things that affect us in the way pleasures and pains do could have value, or why pleasures would be on the value side and pains on the disvalue side.

Of course, there are fundamental differences between presenting an act under a naturalistic guise of <maximizing happiness> vs. presenting it under a normative guise of <what there is most reason to do>. The one has a *directly* deliberative role and *analytic* entailments to claims about how one ought to act that the other lacks. But do we need a *metaphysical* difference in facts or properties to understand this? The question whether to describe these differences in terms of differences in "non-ontological" properties and facts, or differences in non-ontological conceptual roles, seems to me not the sort of thing that could make Parfit's philosophical life efforts in vain, or make Naturalists like Hume or Mill close kin with Nihilists. Or that risks making nothing matter.

Parfit argues that, while natural kind concepts like <water> and <heat> have a "gap" in them, waiting to be filled by some underlying property, evaluative concepts do not (302)—this would seem to count against the conceptual-role story. But it should be recalled that the concept <water> was taken as picking out a *fundamental substance* from Antiquity to the Early Modern period, leaving no "space" for underlying composition by more elementary substances. Similar remarks apply to mentalistic concepts and *mental substance*, or the concept of <vital force> or <life>, where there was no pre-conceived "space" for realization by purely physical systems (cf. 324). And what could be more "undeniably in different categories" than the mental vs. the physical (as Parfit's example of "Experiences are stones" suggests), or colors vs. colorless vibrations, or living beings vs. aggregates of lifeless matter?

It seems better, in a methodological spirit, not to restrict the possibility of finding an underlying explanation or reduction of commonsensical categories by limiting this only to those concepts “ready made” for such explanation or reduction, or to those cases that are intuitively “close”. Certainly history should caution us on this score, and speak against such considerations to dismiss a philosophical approach as inappropriate or incoherent.

In such an open-minded spirit, the concepts, distinctions, practices, etc. of everyday life and prior philosophizing *do* have a “gap to be filled”. From everyday, non-controversial “senses” of things philosophers extract roles or job descriptions, asking: (1) What would it take for this role or job to be filled, or nearly so? And (2) Is there—or could there be—in humans and the world anything that would answer to this? Nor is this a Naturalistic obsession. Kant as much as Hume proceeded from the commonsense understanding of morality to frame questions about what it would take for morality in this sense to be vindicated, and then asked whether a rigorous account of the human mind and the world it inhabits could supply this—so it not turn out that “its secret basis is merely some high-flown fantasticality, and that we may have misunderstood the purpose of nature in attaching reason to our will as its governor” (Kant, 1785, 6:395), and that our “sense of morals ... when reflecting on itself, ... approves of those principles, from when it is deriv’d” (Hume, 1739, III.iii.6). In this sense, they were climbing the same meta-ethical mountain, too, using different gear—“transcendental” vs. “experimental” methods—but aiming for the same summit.<sup>13</sup>

I think the same can fairly be said of Parfit and contemporary Soft Naturalists. Our moral concepts do not have a “gap” ready to be filled by “non-ontological” non-natural properties and facts, either. But Parfit thinks there must be some such properties and facts for normativity to be other than an illusion—and argues that we can indeed give the beginnings of a credible understanding of how there might exist such properties and facts, and how we might have semantic and epistemic access to them, compatibly with accepting the contemporary scientific view of mind, language, and the world. Is this so deeply different from the Soft Naturalist’s approach and ambitions? Might we be scaling the same meta-ethical mountain after all? Meet you at the summit!<sup>14</sup>

---

<sup>13</sup> And with some rather surprising similarities in how they conceived the subjective conditions for normativity, or, as Kant would put it, “the subjective conditions for receptivity to the concept of duty” (see Railton, 1999).

<sup>14</sup> Parfit does suggest that some Analytic Naturalists might end up sharing a mountain with him, were they to “give up [their] Analytic Naturalism” (363)—a rather one-sided reconciliation. Here I am not imagining that Parfit give up his commitment to “non-ontological”, informationally- (or perhaps even linguistically-) individuated, non-natural properties and facts.

As always, I am grateful to colleagues for discussion, and especially to Allan Gibbard and my former colleague Stephen Darwall. Like everyone in the field, I also owe a great deal to Derek Parfit. The errors are of course mine alone.

## References

- Field, Hartry (1993). *Science without Numbers: A Defense of Nominalism*. Princeton: Princeton University Press.
- Hume, David (1739). *Treatise of Human Nature*. Ed. L.A. Selby-Bigge. Oxford: Clarendon, 1888.
- Jackson, Frank (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Clarendon.
- Kant, Immanuel (1785). *Groundwork of the Metaphysics of Morals*. Trans. H.J. Paton. New York: Harper, 1956.
- Lewis, David (1970). "How to Define Theoretical Terms". *Journal of Philosophy* **67**: 427-446.
- Lewis, David (1983). "New Work for a Theory of Universals". *Australasian Journal of Philosophy* **61**: 343-377.
- Lewis, David (1986). *On the Plurality of Worlds*. Oxford: Blackwell.
- Mackie, J.L. (1977). *Ethics: Inventing Right and Wrong*. Hammondsworth: Penguin.
- Railton, Peter (1993). "Noncognitivism about Rationality: Benefits, Costs, and an Alternative". *Philosophical Issues* **4**: 36-51.
- Railton, Peter (1997). "Made in the Shade: Moral Compatibilism and the Aims of Moral Theory". *Canadian Journal of Philosophy*, suppl. vol. **21**: 79-106.
- Railton, Peter (1999). "Normative Force and Normative Freedom: Hume and Kant, but not Hume versus Kant". *Ratio* **12**: 320-353.
- Yablo, Stephen (2001). "Go Figure: A Path through Fictionalism". *Midwest Studies in Philosophy* **25**: 72-102.