

# Finite Beings, Finite Goods: The Semantics, Metaphysics and Ethics of Naturalist Consequentialism, Part I<sup>1</sup>

RICHARD BOYD

*Cornell University*

## 0. Overview.

0.0. *Theistic Ethics as a Challenge and a Diagnostic Tool.* Naturalistic conceptions in metaethics come in many varieties. Many philosophers who have sought to situate moral reasoning in a naturalistic metaphysical conception have thought it necessary to adopt non-cognitivist, prescriptivist, projectivist, relativist, or otherwise deflationary conceptions. Recently there has been a revival of interest in non-deflationary moral realist<sup>2</sup> approaches to ethical naturalism (e.g., Sturgeon 1984; Brink 1984, 1989; Miller 1985; Boyd 1988, 1995). Many non-deflationary approaches have exploited the resources of non-empiricist “causal” or “naturalistic” conceptions of reference and of kind definitions in service of the “naturalistic” metaphilosophical conception that substantive moral questions, and questions about the metaphysics of morals, are broadly *a posteriori* questions, somewhat analogous to scientific questions, and are (in consequence) not amenable to *a priori* resolution by “conceptual analysis.”

Many naturalist moral realists have also advocated some version or other of consequentialism as the substantive naturalist moral theory to which they are committed. Indeed, although nothing like entailment between these positions obtains, the idea that moral questions are questions about how we can

---

<sup>1</sup> This paper began life as a draft of comments on Adams' *Finite and Infinite Goods* for an American Philosophical Association symposium in December 2000. I thank Professor Adams for his response to that draft. I also thank Scott MacDonald and Nicholas Sturgeon for their help with the issues discussed here.

<sup>2</sup> It is commonplace to think of moral realism as encompassing non-debunking conceptions (naturalistic or otherwise) according to which moral statements have (often enough) truth values which are relevantly independent of our particular moral conceptions. Actually, some thoroughly debunking conceptions of the semantics and metaphysics of morals should probably also be thought of as versions of moral realism. See Boyd 1995. Nevertheless, I'll usually follow current fashion in my use of the expression “moral realism.”

help each other flourish seems central to contemporary naturalist moral realism. In a certain sense, some version of consequentialism seems to be the *natural* position for naturalist moral realists.

In the light of this connection, and of the importance of “naturalistic” semantic and metaphilosophical conceptions to contemporary naturalist moral realism, Robert Adams’ *Finite and Infinite Goods* makes an especially important contribution to our understanding of the naturalist realist project in metaethics. In the first place, Adams deploys the resources of a version of “naturalistic” semantics and metaphilosophy in defending the (supernaturalist) conclusion that God is the good and that the goodness of particular things consists in appropriate respects of resemblance to Her (thereby demonstrating the important point that the important point about “naturalistic” semantic and metaphilosophical conceptions is not always that they are *naturalistic*). The details of the particular semantic and metaphilosophical conception Adams deploys are indicative of important issues regarding the semantics of moral realist positions, naturalist and supernaturalist alike.

More importantly, Adams offers a number of criticisms, both of naturalistic versions of moral realism, and of consequentialism. The fact that he shares, with the naturalist moral realists he criticizes, a broadly “naturalistic” conception of the semantics of moral discourse and of metaphilosophy makes his criticisms especially cogent: they are, in essence, “in house” criticisms of particular versions of the sort of moral realist project in which Adams himself is engaged. Indeed, I shall argue, his criticisms allow us to diagnose ways in which the epistemological, semantic, and metaphysical foundations of naturalist moral realism, and of consequentialism, need to be elaborated if the realist naturalist project in metaethics is to succeed. It is the project of the present essay to begin to provide the required elaborations.

In order to appreciate Adams’ criticisms we need an overview of Adams’ own positions.

### 0.1. *Adams on the Good.*

0.1.0. *The Good and the Right.* According to Adams, the good is prior to the right. Things are good in so far as they are, in various respects, excellent, and excellence is a matter of appropriate resemblance to the infinite transcendent good which is God. The primary ethical dictum is that we should love the good. Duties are derivative matters involving social requirements underwritten by divine commands. The notion of loving the good is extensively filled out with subtle and morally sensitive discussions of the notion of vocation and of *symbolic* as well as instrumental ways of exhibiting love for the good. A key notion here is that loving the good is a matter of being for the good: of *alliance* with the good (=God) which is expressed, among other ways, through appropriate alliances with others.

0.1.1. *Excellence (and thus the Good) as Non-instrumental.* Adams insists that the good is to be understood intrinsically rather than instrumentally. He makes it clear in the Introduction that one of his central targets is the version of Christian ethics with which he grew up "...the well-being of persons and the quality of personal relationships being its primary (and it sometimes seemed only) concerns (4)."

This theme reappears in Chapter 7 where Adams indicates that he is concerned to reject ethical conceptions so focused on the improvement of society or the well-being of others that they made it hard for him "to see how it could be ethically good to be interested in beauty and truth, art or philosophy, for their own sakes, since such interests were not interests in persons (181)," and where he addresses the question of whether "the theistic ideal must denigrate aesthetic and intellectual values, for example, in the name of narrowly moral concerns (185)."

0.2. *The Unity of the Good and Adams' Critique of Consequentialism.* Closely related to his insistence that the excellences that are aspects of the good are intrinsically rather than just instrumentally good, is Adams' critique of consequentialist conceptions in ethics. Consequentialism is something of a moving target in *Finite and Infinite Goods*. In the early chapters it appears that all versions are under attack. In Chapter 13 ("Vocation") however important concessions seem to be made to approaches which fall under the category "indirect consequentialism." Despite these complexities, we can, I think, identify two important dimensions to Adams' critiques of consequentialism, each of them raising doubts about the ability of the consequentialist to provide an account of the ways in which morally relevant goods can provide a unified standard of value. [I here ignore distinctly theological difficulties of the sort he discusses in Chapter 9 ("Symbolic Value") as lying beyond the scope of the present project.]

0.2.0. *Consequentialism and the Fragmentation of the Good.* A central element in Adams' critique concerns the issue discussed above of the ethical place of artistic and intellectual values, which are not in obvious ways matters of valuing the well-being of others. Adams thinks that the consequentialist will have a hard time according these values the ethical standing which they obviously have. According to him, an advantage of his conception over ethical conceptions grounded in consequentialism is that it avoids a fragmented conception of the good. If one has a consequentialist conception of ethical goodness, with an emphasis on issues about the welfare of persons, then one will have to see moral goods (so understood) as competing (with respect to motivation and with respect to rational choice) with *non-moral* aesthetic, intellectual (and other) goods. Such a conception Adams sees as

neither psychologically desirable nor metaphysically accurate. It underestimates the unity of the good in a way that the proposed excellence-based theistic conception does not.

0.2.1. *Non-Consequential Moral Considerations: Alliance and Symbolic Value.* A further consideration leads Adams to a non-consequentialist conclusion. In Chapter 9 ("Symbolic Value") he considers "hopeless" cases in which—with respect to some morally significant situation (even in the narrowest sense "morally significant") the right (or morally preferable) thing to do seems to have no good consequences to which a consequentialist could appeal. Such cases include those in which symbolic opposition to tyranny seems the right action to take even though it can be expected to have no effect on the political situation or the conditions of the oppressed.

Adams understands the appropriateness of symbolic opposition in such cases in terms of the notion mentioned earlier of being (symbolically, if in no other way) *for the good*, especially by being (symbolically at least) for those who suffer. He complains that many modern ethical theories—not just utilitarian ones—fail to provide guidance in situations of total or partial helplessness. They "construe the task of ethics too narrowly, as guidance for *action* (224; emphasis his)." I assume that this criticism is supposed to be directed not just at utilitarian conceptions but at consequentialist conceptions more generally with, perhaps, the exception of the "indirect consequentialism" mentioned later in Chapter 13 ("Vocation"). Here again, the issue is one of the unity of the good since, according to Adams, the consequentialist will, quite implausibly, have to treat standards for behavior in helpless situations as distinct from standards of moral goodness.

0.3. *Adams on Moral Semantics and the Limitations of Naturalism.* Adams quite consciously models his treatment of the semantics of moral categories on recent naturalistic developments in the semantics of scientific terms. At the same time, he identifies what he sees as weaknesses in any thoroughly naturalistic approach to the semantics of ethical terms. In Chapter 2 ("The Transcendence of the Good") he articulates his version of the "open question" argument, maintaining that a naturalistic treatment of the semantics of "good" would fail to account for the possibility of a *critical stance* which is fundamentally important in ethics.

In the next chapter ("Well-Being and Excellence") Adams criticizes as well accounts which characterize a person's (ethically relevant) well-being in terms of counterfactuals regarding idealizations of her actual desires and preference structures. Of course, one need not be a naturalist to subscribe to such a counterfactual account. Still, as Adams notes, part of the motivation for such accounts has often been the reduction of the notion of a person's good to

empirical (and presumably natural) facts. So it seems reasonable to take his critique of them as contributing to Adams' case against naturalistic conceptions in ethics.

0.4. *Adams on the Referential Semantics of "Good": "Naturalistic" Semantics without Naturalism.* Adams holds that the sort of goodness that is ethically relevant is *excellence*, understood as a certain sort of *resemblance to God*. This claim is not analytic: it is not an analysis of the meaning of "good." Adams defends it instead as a claim about the *nature* of the good analogous to the claim that the nature of water is  $H_2O$ . He correctly asserts that the results of recent naturalistic work on the semantics of scientific terms—work which provides a semantic underpinning for the view that the essence of water is  $H_2O$  even though this is not a consequence of any analytic definition of the term "water"—can be extended to categories other than natural kinds, and he proposes such an extension in the case of moral categories.

Adams makes a particular proposal about the semantics of words like "water" and "good," which lack analytic definitions—a proposal about the natures of their referents. He says "What is given by the meaning, or perhaps more broadly by the use of words, is a role that the nature is to play. If there is a single candidate that best fills the role, that will be the nature of the thing" (16). The equivocation between meaning and use in this formulation will be important later. For now, we should note that Adams seems mainly to focus on meaning rather than use. Most of the time he seems to hold (as does, for example, Sidelle 1989; note however that Adams' account of modality (46) is richer than Sidelle's) that it is an analytic (or at least a "conceptual") truth that the referent of such a word, *if such a referent exists*, has the properties associated with the conceptually central aspects of our use of the word. Thus, in discussing internalism Adams acknowledges "...as a conceptual truth that if anything is good, in the sense of 'excellent,' it is *good* for us to love it, admire it, and want to be related to it, whether we do in fact or not" (25). On such a view what is not analytic is that a referent of the required sort exists, or what metaphysical nature underwrites its playing the analytically (or "conceptually") specified role, if it does.

In fact, I doubt that there is any analytic or conceptual connection between excellence and rationality of the sort Adams suggests, in part because I doubt that there is any *conceptually* or *analytically* specifiable concept of rationality for any other concept to be conceptually related to.<sup>3</sup> For present purposes, however, I want to focus on the way in which the reference-fixing role asso-

---

<sup>3</sup> I discuss the nature of inductive rationality in Boyd 1992 and conclude that the nature of inductive rationality is given a *a posteriori* homeostatic property cluster definition of the sort discussed later in this essay. I think that a successful account of the rationality of preferences, choices and actions would attribute to that sort of rationality a similarly a *a posteriori* definition.

ciated with a word is determined. Understanding that will hold, I believe, one of the keys to assessing Adams' critique of ethical naturalism, to which I now turn.

## 1. Adams' Critique of Ethical Naturalism.

1.0. *The Critical Stance.* In Chapter 2 ("The Transcendence of the Good") Adams offers a spirited defense of the ethical realism, to which he and ethical naturalists like me both subscribe, against a variety of epistemological challenges. He then offers a very interesting critique of the sort of naturalistic ethical realism which I and others have recently been defending. This critique is his version of Moore's "open question" argument.

Here's the argument. Adams envisions the ethical naturalist's dream: the scientific confirmation of a particular naturalistic account of the ethical uses of the term "good." He imagines that it is discovered that a particular family of natural properties causally regulates those uses of "good" in the way required by the naturalist's semantic theory, so that—by the naturalist's standards—it is scientifically confirmed that that family of properties constitutes the nature of the good.

[In fact, what Adams imagines is the confirmation, in this sense, of my homeostatic property cluster theory of the good (Boyd 1988), which he takes to be a form of act consequentialism. As it happens, I intended it as a version of the doctrine which he calls "indirect consequentialism," but that is irrelevant to the topic at hand, since Adams' critique is supposed to work against any form of naturalism.]

Adams maintains that this picture of how a naturalistic conception of the good might be scientifically confirmed fails to account for a *critical stance* which is a crucial part of the "intentional framework" within which we use ethical terms. He maintains that, even supposing that some naturalistic conception had been "confirmed" in this way according to naturalists' evidential standards, it would *always* be possible to cogently raise the question of whether or not the relevant family of natural properties is *really* (the) good. As he sees it, no naturalistic conception can do justice to this fact, because it follows from such a conception that if we did know all of the relevant natural facts, then there would be no rational room for such a question. The naturalist's metaphysics of morals is thus incompatible with an important fact of moral epistemology.

By contrast, according to Adams, a theistic metaphysical foundation rationalizes the relevant epistemological principle. Because God is not fully comprehensible by any finite mind, a metaphysical theory which identifies goodness with God, and particular excellences with resemblance to Her, underwrites the judgment that no amount of scientific knowledge could finally settle issues in moral theory.

Adams considers one possible response which someone with my views might offer. I think that ethical goodness is probably defined by what I call a homeostatic property cluster: a family of properties of actions, policies, character traits and the like which are aspects of, or contribute to, human flourishing and which are such that they exhibit a sort of homeostatic causal unity: under suitable conditions their instantiations are (causally) mutually supporting. This homeostatic causal unity is the only sort of unity which, according to the view I defend, the good possesses. I maintain that it is a central concern of social and political morality to identify and implement social and economic structures and policies which enhance this homeostatic unity, so that increasingly deep manifestations of human flourishing will be able to co-exist in a mutually sustaining way.

In discussing alternatives to homeostatic consequentialist moral realism (see Boyd 1988 and especially Boyd 1995) I considered the possibility of a certain limited sort of moral relativism. It seemed to me that, with respect to the project of enhancing the homeostatic unity of such human goods as we can collectively possess, it might turn out that there are two (or more, but let's make it two for convenience) paths of moral development, each in a direction plainly preferable to our current situation, but such that their endpoints would involve morally incommensurable ways of achieving homeostatic unity of features of human well being, and such that there exists no alternative path to an arrangement which relevantly "splits the difference" between them. One possible source of such incommensurability, I suggested, might be that, because people's psychological and social needs are to a significant extent themselves socially determined, the nature of flourishing might differ significantly for people who grew up under the conditions achieved in the two different developmental endpoints.

I suggested that the hypothesis that this is the actual state of affairs with respect to the possibilities of moral progress might serve as a naturalistic formulation of the most (perhaps the only) plausible version of moral relativism. If this hypothesis is true, I suggested, it would be appropriate to treat the term "good" (in its current ethical uses) as *partially denoting* each of the two different version of (homeostatically unified) human flourishing.

Such a situation, if it obtained, would refute moral realism as that doctrine is ordinarily understood. Nevertheless, I argued, the situation I envisioned would sustain a basically realist conception of moral and ethical discourse since moral discourse would turn out to be about two closely related families of natural phenomena (I have in mind, goodness, justice, duty,...etc. as they would be manifest in the arrangements instantiated in the two different endpoints) and since no antirealist retreat to conventionalism or noncognitivism would be involved.

Adams considers the possibility that a naturalistic moral realist would invoke something like this sort of partial denotation in order to account for the critical stance appropriate to moral reasoning. As I understand it, what he has in mind is that—in the situation in which act consequentialism (or some other naturalist moral theory) is confirmed by *naturalist standards* but in which Adams (for example) dissents and favors a non-consequentialist theist account—the naturalist might hold that the term “good” partially denotes each of two phenomena: the naturalistic “good” identified by the evidence for act consequentialism, and some other intrinsic “good” answering to the theist’s concerns. He concludes that, since this approach would force the naturalist to be “a nonrealist” about the question of consequentialism, it is unattractive.

Instead, he proposes the theistic metaphysical explanation for the cogency of the critical stance: that the transcendent good (=God) lies beyond complete human comprehension.

Although I don’t think that the naturalists’ best response to the challenge Adams poses involves recourse to partial denotation in the way indicated, I do think that Adams’ discussion of partial denotation raises really important questions, not just about the semantics of ethical naturalism in particular, but about naturalistic semantic theory in general. My response to these questions will be somewhat controversial, so I propose first to argue that, whatever subtle difficulties Adams has raised for naturalistic semantics, the critical stance itself does not pose a serious challenge to ethical naturalism. I’ll thus begin my discussion by raising several points in defense of this last claim.

1.1. *The Critical Stance and Reference-Fixing Roles.* I indicated earlier that Adams holds that the (nature of the) referent of a word whose semantics is like that of natural kind terms is determined by a *role* which is itself “... given by the meaning, or perhaps more broadly by the use” of the word in question, so that “(i)f there is a single candidate that best fills the role, that will be the nature of the thing.” I mentioned that Adams ordinarily seems to think of this role as being accessible to conceptual analysis so that it is analytic (or something like that) that the referent of a word, if it has a referent, will be something of which the properties indicated by the word’s *meaning* are true.

Of course, the formula Adams offers for specifying natures does not require such an interpretation. In the first place, of course, it leaves open the possibility that the nature in question is determined by features of the word’s use other than those readily accessible to conceptual analysis. Moreover, however the role is determined, Adams’ formula requires only that the nature in question be “the single candidate which best fills the role,” so that that nature need not be such as to fit the role *exactly*.



In particular, even if we conclude that it is part of our conception of the good that, no matter what the empirical evidence, general ethical claims are always “open questions,” as Adams’ argument requires (more on this later), the option remains for the ethical naturalist to hold that this is merely an *approximately right* judgment about the epistemology of ethical generalizations. Indeed, Adams raises this possibility himself, suggesting that the naturalist might endorse the critical stance as appropriate in the light of the uncertainties of empirical investigations. This approach would not, he observes, deliver a justification for the critical stance under the assumption that all possible empirical evidence had been properly assessed.

I’ll have more to say about that (very highly) counterfactual possibility later on. For now, let’s note that—given Adams’ own conception of how the natures corresponding to words are determined—the nature of the good need not *exactly* underwrite the aspects of the critical stance which he sees as part of our conception of the good. Only if one adopts a conception of reference which ill suits Adams’ broader project of defending a *non-analytic* conception of the good would one think otherwise.

1.2. *The Critical Stance, Critically Examined.* I want to turn now to the question of the critical stance itself. As Adams anticipates, I agree that the epistemically correct approach to ethical theories is *something like* the critical stance he articulates. I think that it is a good methodological principle for the ethical naturalist (or anyone else for that matter) to be generally skeptical about taking empirical evidence to be definitive with respect to general issues in ethical theory. As Adams anticipates, I would defend this generally critical attitude by pointing to the complexity of the empirical issues which are involved in serious ethical disputes.

Of course, Adams wants to defend more than a generally critical attitude towards ethical theories. He proposes that we should treat any ethical theory as still an open question even when faced with empirical evidence which would, by an ethical naturalist’s standards, confirm it. There are several points of philosophical method which we need to get straight before we investigate this proposal.

1.2.0. *Methodological Clarifications.* In the first place, for reasons I’ll discuss later, the sort of critical stance Adams advocates is much easier to defend from a theistic conception of ethics than from most others. I take it that Adams does not intend for his criticism of ethical naturalism to depend on any theistic hypothesis. Instead, as I understand it, he intends to pose a dilemma for the naturalist who is also an ethical realist: The critical stance (understood as Adams understands it) is pretty obviously the appropriate epistemic attitude. A naturalistic conception can explain its appropriateness only

by abandoning ethical realism. Thus, the naturalistic ethical realist must choose between her ethical realism and her naturalism.

Adams suggests giving up naturalism.

In the light of this reconstruction of Adams' argument, I'll try to assess his version of the critical stance without presupposing any particular answer to the question of theism or to any similarly esoteric metaphysical question.

Secondly, it is important to note that the plausibility of Adams' criticism of naturalist moral realism depends on his being concerned to defend a non-trivial version of the critical stance. What I have in mind is this. According to the naturalistic ethical realist, the epistemology of moral reasoning is basically like that of everyday and scientific reasoning about matters of natural fact. Whatever the right account of such reasoning is, it will certainly underwrite the Humean principle that scientific and everyday factual generalizations do not follow deductively from the empirical evidence which supports them. The naturalist ethical realist will certainly have the resources to argue that the same holds true for ethical generalizations.

Thus, in one sense of "open question," she will be able to consistently accept the claim that any ethical generalization remains an open question given the empirical evidence which by her own standards confirms it. This critical-stance-on-the-cheap is not philosophically unimportant—it addresses Moore's own open question argument for example—but it is not a response to the interesting challenge Adams poses. He maintains that general ethical claims remain open questions in the light of empirical evidence, even when that evidence is assessed by rational evidential standards which go beyond the principles of deductive logic.

Finally it is important to see, as Adams clearly does, that his critique of naturalism is not supposed to rest on his choice of act consequentialism as a sample naturalist ethical theory. The same sort of critical stance should be appropriate for any naturalistic theory, given any body of empirical evidence which, by naturalist standards, would seem to confirm it.

With all these methodological reminders in place, I want to argue that it is by no means obvious that the critical stance, as Adams portrays it, *is* part of our current conception of ethical reasoning. I think that choosing act consequentialism as the sample naturalist moral theory has the effect of distorting the epistemological judgments upon which the plausibility of (Adams' version of) the critical stance depends.

The problem with the example, I suggest, is the overwhelming implausibility, from a naturalistic point of view, of any version of act consequentialism. First, let me explain why naturalistic act consequentialism is so profoundly implausible. I'll then explain why its implausibility poses a

methodological problem when it is deployed as a sample naturalistic ethical position.

1.2.1. *The Theoretical Implausibility of Act Consequentialism.* Adams provides just a sketch of a theory of reference for terms which are like natural kind terms in having no *a priori* definitions, and I'll suggest an alternative more complete theory later on. Nevertheless, there are, I believe, two principles about reference of such terms with which almost any plausible "naturalistic" semantic theory will agree. In the first place, our use of such a term must afford us "epistemic access" to its referent. If a term, *t*, of this sort refers to a phenomenon, *p*, then the real properties of *p* (or of *p*'s, depending on the sort of phenomenon *p* is) must contribute to the regulation of our *t*-beliefs in such a way that, under relevant circumstances, many of the things we predicate of *t* tend systematically to be approximately true of *p* (Boyd 1983, 1989, 1993, 1999c, 2001b). Note that I do not mean to require that the most fundamental properties of *p* (or *p*'s) must reliably regulate our meta-physical *t*-belief.

Depending on the term and the contexts of its use, the relevant epistemic capacity may reside in users of the term generally, or in experts to whom they defer. Depending on the context, the extent of the required deference may vary. It may even happen that the relevant epistemic access is achieved despite, rather than because of, the efforts of experts. [Arguably this is true with respect to terms like "intelligence," "aptitude," "just war," "human nature," "race," "social class," "sex differences," and "altruism," which are central to disciplines very much under the influence of social ideology.]

In any event, however subtle the differences between these various sorts of cases, *t* cannot refer to *p* unless there are some people who, under ordinary circumstances, are at least pretty good at finding out about *p* and who reflect this capacity in what they say using *t*.

In addition to this *epistemic access condition*, we should also acknowledge an *achievement explanation condition*: in order for *t* to refer to *p*, the epistemic access which uses of *t* affords speakers to the real properties of *p* must (help to) *explain* the theoretical and/or practical successes achieved in the domains of inquiry or of practice to which *t*-talk is central. This condition reflects the basic philosophical motivation for acknowledging *a posteriori* definitions for natural (and other) kinds in the first place: to explain, as empiricist conceptions of language and of classification cannot, how the non-"nominal" uses of scientific terms, and the associated "metaphysical" classificatory practices, contribute to the inductive and explanatory success of science.

Conditions like these are probably tacitly at work when philosophers formulate or apply best-role-satisfier accounts of reference like Adams', espe-

cially in cases (like those of *natural* kinds) in which the roles are causal-explanatory. In any event, they are important in ruling out some implausible candidates for the reference relation which might otherwise not be excluded.

Consider, for example, cases in which a natural kind term, *t*, is associated with a very abstractly formulated causal role specification, *S*, and in which there are phenomena, *p*<sub>1</sub> and *p*<sub>2</sub> such that *p*<sub>1</sub> more closely satisfies *S* than does *p*<sub>2</sub>, but practices among *t*-users affords them significant epistemic access to *p*<sub>2</sub> but not to *p*<sub>1</sub>. The whole point of the "naturalistic" critique of purely descriptivist conceptions of reference is that in such cases there is a *prima facie* reason to assign *p*<sub>2</sub> rather than *p*<sub>1</sub> as the referent of *t*. "Naturalistic" conceptions of reference are reasonably called "causal" conceptions largely because the epistemic access relation between a term and its referent is (at least for proper names and natural kind terms) ordinarily underwritten by epistemically relevant causal connections between the use of the term in the relevant community and (instances of) its referent.

The achievement explanation condition is also tacitly presupposed in standard applications of naturalistic conceptions of reference. Consider the word "water" and the (correct) idea that it refers, and has referred for a long time, to the natural kind whose definition is provided by the formula  $H_2O$ . Now, almost all of the samples to which the term "water" (unmodified) has been applied have in fact been samples of very dilute carbonic acid, owing to the presence of  $CO_2$  in the atmosphere and its solubility in water. Thus, at least for a long historical span, the use of the term "water" afforded speakers better epistemic access to dilute carbonic acid than it did to pure  $H_2O$ . In assigning  $H_2O$  rather than dilute carbonic acid as the referent of "water" (as we should) we are (tacitly or explicitly) recognizing the importance of the achievement explanation condition: it is the connection between uses of "water" and  $H_2O$  molecules, and not the (epistemically even closer) connection between those uses and samples of dilute carbonic acid, which explains how our use of "water" contributed (and contributes) to our inductive, explanatory and practical successes.

Consider now the implications of these two conditions for act consequentialism. Suppose that the term "good" refers to (something like) the property which a particular (token) action has if, among the actions available to the actor, it is among those which would contribute most to human flourishing, as act consequentialism requires. In order for this to be so, it would have to be the case that (a) our use of the term "good" (in moral contexts) affords us epistemic access to that property, and (b) the extent and reliability of that epistemic access helps to explain our successes—such as they are—in matters moral. Now of course it is a controversial issue just what sorts of successes would count as *moral* successes. Let us suppose—as any consequentialist must—that the successes in question are matters of our being able to contrib-

ute to human flourishing. On that assumption, what the *act* consequentialism considered by Adams entails is that

(a) we are, often enough, pretty good, in practice, at applying the word “good” (in contexts of moral deliberation) to those choices which have the property in question (even if we don’t know that it defines moral goodness),

(b) good enough, in fact, that our success in tracking the property in question with our moral uses of the word “good” is central to the explanation of our achievements in so far as contributing to human flourishing is concerned.

The supposition that this is so is profoundly implausible theoretically, especially for someone with a purely naturalistic conception of human capacities. We lack both the information gathering and the information processing capacity to anticipate with any reliability the consequence for human flourishing of particular action choices *or even to retrospectively assess the overall consequences for flourishing of particular actions*. It is in all probability physically impossible for systems with our information gathering and computational structures to exhibit the required capacity. Instead, in so far as anything like the act consequentialist hypothesis represented by (a) and (b) is true, it would have to be some version of what Adams apparently calls “indirect consequentialism,” which I take to be related to act consequentialism in the way rule utilitarianism is related to act utilitarianism.

I want at this point to consider a plausible objection to my argument against act consequentialism. Someone might reasonably object that the considerations I have rehearsed fail to take into account the appropriate role of idealization in establishing the referential semantics of moral terms. She might argue as follows.

According to the consequentialist, the aim of moral practice is to guide our actions so as to enhance human flourishing. In consequence (or perhaps, in consequence of this aim properly understood) the following counterfactual is true: If we could figure out which of our actions satisfy the *act* consequentialist’s proposed definition of the good, those would be the right actions to perform. Because actions conforming to act consequentialist standards would be the good ones to perform under such idealized conditions, we should accept the act consequentialist conception of the nature of the good, and treat the indirect consequentialist’s proposal as a *practical* proposal for *implementing* consequentialist morality rather than as providing an accurate definition of the good.<sup>4</sup>

There is much to say about this sort of argument (especially about the doubtful intelligibility of the profoundly counter-legal counterfactual at its heart), and much to say about the role of idealization, in moral practice and in

---

<sup>4</sup> I thank Nicholas Sturgeon for emphasizing the importance of this objection.

semantic theory. For present purposes, however, it suffices to note that—on a “naturalistic” conception of essences and of the semantics of “natural” kind terms, of the sort Adam’s himself accepts with respect to the metaphysics of morals—idealizations in the theory of kind definitions are subject to the epistemic access and achievement explanation conditions. That the proposal to accept act consequentialism as providing the natural definition of the good violates these constraints profoundly can be seen by examining an analogous case.

The aim of preventive medicine is, we may suppose, to prolong healthy and satisfying lives.<sup>5</sup> The term “healthy” as it applies to diets is, let us presume, a natural kind term in the discourse of preventive medicine. A reasonable naturalistic proposal about the essence of healthiness of diets is that it is defined (*a posteriori* of course) by some complex family of health related properties of the sort which careful investigation might indicate as contributing to, e.g., the prevention of heart disease, cancers, and the like.

Suppose that someone were to object to this semantic conception by arguing that, if it were possible to preserve satisfying life indefinitely, it would be good practice of preventive medicine to do so, and that *therefore* the real definition of “healthy,” when applied to diets in preventive medicine contexts, is the property of *tending towards* the indefinite perpetuation of satisfying life. She might suggest that the more modest definitional proposal bears to her idealized proposal just the relation which, according to the criticism we are considering, the indirect consequentialists proposal bears to the act consequentialist proposal.

She would be right about the analogy between the two cases, but certainly wrong about healthiness, precisely because her proposal represents an idealization which violates the epistemic access and achievement conditions. In some extraordinary sense of approximation, our dietary practices in preventive medicine might be said to approximate those practices which might—in some very distant possible worlds—preserve life indefinitely, but this sort of approximation is irrelevant to the explanation of our achievements in preventive medicine even if good preventive medicine in such worlds would aim to preserve life indefinitely.

Of course, the same is true with respect to the proposed idealized “natural” definition of the good, on any conception of natural definitions compatible either with Adams’ project or with the naturalist moral realists projects he criticizes. The fact, if it is a fact, that our moral aim of contributing to human flourishing would—in some extremely distant, nomologically impossible, possible world—be best served by acting on act consequentialist calculations is irrelevant to the question of the natural definition of the good.

---

<sup>5</sup> This is not the place to discuss the alternative proposal that the purpose of preventive medicine is really to save employers and insurance companies money.

Moral: If you are going to be a naturalist consequentialist, be an “indirect” consequentialist. The standard philosophical arguments which favor rule utilitarianism over act utilitarianism also show that defending naturalistic act consequentialism requires positing human information acquisition and processing mechanism which make no sense in terms of our best scientific understanding of the natural world.

Thus when we consult our epistemic judgments (our “intuitions,” to use the term preferred by many philosophers) about what our rational epistemological options would be if there were (what would be by the ethical naturalist’s standards) empirical evidence confirming that “good” refers to the act consequentialist’s favorite candidate, we are engaging in an exercise somewhat akin to asking what our rational epistemic options would be if the best available empirical evidence confirmed the hypothesis that frogs fairly reliably computed the orbits of the inner planets. I think that neither exercise provides much philosophical insight about the epistemology of ethics or (in the second case) herpetology.

1.2.2. *Interlude: The Philosophy of Science, Theory Dependence, and Quasi-analyticity.* I want to defend this latter claim by exploiting some insights regarding the methodology of the empirical sciences which Adams himself emphasizes. The first of these concerns the profound theory dependence of the methods by which scientific questions—including questions about the natures of things—are properly answered. In Chapter One Adams considers objections to the effect that his treatment of the good is circular, since in articulating and justifying his conception of the good he employs lots of related normative notions like, for example, the notion of “excellence” and of a phenomenon’s being the “best fit” to the role associated with an ethical term.

His response is entirely appropriate. He reminds the reader that he is not trying to introduce the terms “good,” “excellent,” or any other normative terms, but to figure out what the natures of goodness, excellence and other normative phenomena are. It is no more an objection to his approach that he employs these notions than it would be an objection to the approach of the chemist who claims that  $\text{water} = \text{H}_2\text{O}$  that she relies on lots of information about water and about chemistry in general. In each case, we properly make use of (what we take to be) our best available knowledge about the relevant phenomena in order to try to figure out what their natures are. Inquiry into the natures of things is—in ethics and in science—*theory dependent*.

In the second chapter Adams makes methodological points which depend on further facts about the ways in which scientific and ethical inquiry are theory dependent. This is the chapter in which Adams defends his version of the critical stance, and he prepares the ground by (correctly) indicating, in various ways, that the sorts of scientific considerations on which, according

to naturalists, ethical reasoning depends do not always “trump” moral and ethical judgments. He insists that some moral beliefs are more secure epistemologically than even quite well confirmed scientific theories, arguing, for example, that “...it would surely be crazier to give up the belief that it would be wicked to torture children than to say that there must be something wrong with quantum mechanics (76).”

Later Adams insists, also correctly, that, just as moral realists acknowledge that it is sometimes rational to infer an “ought” from an “is,” the opposite is also true: we are sometimes rationally justified in reaching conclusions regarding empirical factual matters from moral premises.

Each of these points is designed to support Adams’ version of the critical stance by undermining any conception that scientific evidence always trumps moral considerations. Importantly, however, each depends upon methodological principles regarding the theory dependent assessment of evidence, in both ethics and science, which will prove significant in underwriting a critique of Adams’ argument for the critical stance.

Consider first his claim that there are moral principles which are epistemically more securely established than even well established scientific theories. The cogency of this point is a reflection of an important fact about the way in which background theories rationally inform our epistemological judgments: even among beliefs which we accept on good evidence, some are such that we rationally assign to them a much more secure epistemic position than we assign to others.

Indeed, many fundamental scientific laws (as well as some scientific truisms) and many fundamental moral principles have the property which we might call *quasi-analyticity* (see, e.g., Putnam 1962). Because of their conceptual and methodological centrality, even when we know that their justification is *a posteriori* rather than *a priori*, we find it extremely difficult to envision circumstances under which they would be disconfirmed. For as long as they occupy so central a conceptual and methodological role, they are immune from empirical revision, and principles incompatible with them are ineligible for empirical confirmation (let’s call them *quasi-analytically ineligible*). As Putnam indicates, quasi-analyticity and quasi-analytic ineligibility can be altered only by pretty serious conceptual and theoretical “revolutions,” whose directions are all but impossible to anticipate prior to the innovations or crises which precipitate them. The principle that torturing children is wicked and the fundamental laws of quantum mechanics are both candidates for quasi-analyticity.

Consider next Adams’ claim that we may sometimes justifiably infer factual claims from moral premises. This is so because our concepts (empirical and moral) are related through a complex web of mutually ratifying inferential connections so that, in general and not just in ethical discourse,



the adoption of a premise from one domain of inquiry will have methodological ramifications in lots of other domains (see Quine and Ullian 1978).

Adams' (successful) efforts to show that some moral conclusions are relatively immune from scientific refutation depend on the way in which these two features of inferential rationality interact. At a time when a principle (moral, scientific, or whatever) is one of the most epistemically secure—is quasi-analytic or something like that—it is, at that time, largely immune from rational disconfirmation, even though it is not knowable *a priori*, and even though its position in the web of inferential connections may entail that its immunity has significant methodological implications regarding the assessment of principles in other domains.

Before we explore Adams' choice of act consequentialism as a sample naturalistic theory in the light of these facts about theory dependent methods, we need to examine one of their important corollaries. Philosophers often deploy their epistemic judgments (or "intuitions") regarding actual or possible evidential situations as though they were (perhaps imperfect) reflections of *a priori* justifiable inferential principles. The theory dependence of such principles shows that, at least for the case of scientific inferences (and ethical inferences too, if realist ethical naturalism—or Adams' version of realist ethical supernaturalism—is true), that is not the case. Inferential principles, in so far as they are justifiable, are justifiable only *a posteriori* (see Boyd 1989, 1991, 1992, 1999c).

In particular, epistemic intuitions are matters of trained judgement, reflecting explicit or tacit beliefs of the judger or of the community into which she is acculturated. Someone in the fortunate position of having her epistemic judgments rest on approximately true beliefs (about the actual world), will have epistemic "intuitions" which are appropriate to the actual world and to very similar ("nearby") "possible worlds." There is no reason to suppose, and every reason to doubt, that her intuitions will be a reliable guide to the inferences which would be reliable in very different "possible worlds."

In the light of these considerations, I want to examine Adams' question of whether or not it would be rational to accept act consequentialism if the empirical evidence confirmed that doctrine according to the naturalist's evidential standards. Adams consults his epistemic judgment, concludes that the answer is "no," invites us to reach the same conclusion, and takes the result to support his version of the critical stance. I'll argue that, in the light of the quasi-analytic ineligibility of act consequentialism, Adams' exercise tells us little about the epistemology of morals. As a prologue to my argument, I propose an exploration into the epistemology of herpetology.

1.2.3. *Frogs, and an Ambiguity in Counterfactual Epistemological Questions.* Let T be some biological theory which entails that certain frogs compute the orbits of the inner planets; T might, for example, be a theory about frog navigation, or about the timing of breeding in the species in question. Suppose that a philosopher, let's call her Eve, raises the question, "Would one be rationally required to accept T in the light of apparently confirming empirical evidence?" Now T is quasi-analytically ineligible. I'll argue that, as a result, (1) Eve's question is ambiguous, (2) a failure to appreciate the points about theory dependence just rehearsed would lead someone who relied on her epistemic intuitions about T to fail to appreciate the ambiguity in question, and (3) the way in which Eve's question is ambiguous limits the methodological value of appeals to epistemic intuitions, even for someone who appreciates that ambiguity.

Ambiguity first. For an everyday theory, E, which is not quasi-analytically ineligible, the question "Would it be rational to accept E in the light of apparently confirming empirical evidence?" directs our attention to situations which differ from our actual situation (if at all) in being such that the empirical evidence in those situations supports E *by the methodological standards which currently prevail in our actual situation*. Given the peculiar situation of T, however, Eve could reasonably be understood to have asked about the situation of T under two quite different sorts of counterfactual situations. Her question, if made precise, could be either of the following:

(a) Suppose that there were empirical findings about the frogs such that, if facts of that sort were discovered about organisms regarding which an hypothesis like T were not quasi-analytically ineligible, they would confirm about those organisms the hypothesis which T asserts about frogs. Would one then be rationally required to accept T? [Call this the *nearby worlds question about T*.]

(b) Suppose that there were empirical findings which rationally occasioned a theoretical and conceptual "revolution" regarding the perceptual and computational capacities of frogs, in the light of which T was no longer quasi-analytically ineligible, and which confirmed T *by the theory dependent scientific standards underwritten by the new background theoretical framework*. Would one then be rationally required to accept T? [Call this the *distant worlds question about T*.]

At least by naturalist standards, the answer to the nearby worlds question is "no," whereas the answer to the distant worlds question is "yes." Suppose now that Eve has intuitively appreciated the fact that T is quasi-analytically ineligible, but that she has not attributed T's immunity from confirmation to the theory dependence of inferential methods and to the derivative phenome-

non of quasi-analyticity. She might well fail to discern the ambiguity we have just discussed and instead believe that she had discovered a peculiar immunity of T to decisive empirical determination: a *critical sprawl* appropriate to questions in herpetology. [Amphibians, crocodilians and lizards, unlike dinosaurs, mammals and birds, have a *sprawling* posture rather than an upright *stance*.]

There is an additional methodological point to be made about Eve's appeals to her own (or our) epistemological intuitions regarding T. Suppose that Eve does appreciate the ambiguity we have been discussing and that she is concerned to address the distant worlds question, the one about the epistemological fate of T in the aftermath of a scientific revolution. Suppose that Eve is not a naturalist about matters biological; perhaps she is some sort of theist vitalist. For her, perhaps, the affirmative answer to the distant worlds question is not obvious; perhaps her vitalism makes her critical of scientific standards of evidence.

*Even so*, if Eve understands that rational inferences regarding matters of fact have only a *posteriori* justification, she will recognize that she is not justified in taking her unfavorable intuitive epistemological judgments regarding T as applicable under the circumstances envisioned in the distant worlds question. As we have seen, epistemological intuitions, even the best justified ones, are ordinarily suited only to possible worlds very nearby the actual world, and worlds of the sort contemplated in the distant worlds question are, well, quite *distant* from that world.

1.2.4. *The Critical Stance Examined in the Light of a Better Example.* Of course I want to suggest, on behalf of ethical naturalism, that the eminently plausible epistemological intuition that one would not be rationally required to accept act consequentialism even if empirical data appeared to confirm it is a reflection of the fact that act consequentialism is quasi-analytically ineligible for confirmation. To see the strength and plausibility of the relevant epistemological judgment as instead indicating something epistemically special about ethical judgments would, I think, be to make (a much more plausible version of) just the mistake that Eve would make if she concluded that the epistemology of herpetological judgments involves a special critical sprawl.

I am suggesting that our adverse epistemological intuitions about act consequentialism dictate a negative answer to the nearby worlds question about it, but (as they must always) leave the distant worlds question unaddressed. I haven't tried to show that there are no other considerations available, say to a divine command realist, which might establish to her satisfaction a negative answer to the distant worlds question about act consequentialism. What I do contend is that, in so far as they are supposed to persuade us that naturalist

realist metaethical conceptions cannot account for the epistemological peculiarities of act consequentialism, Adams' arguments are unsuccessful

If one really wants to consult counterfactual epistemological judgments in order to see whether or not the critical stance is part of our current conception, the correct approach would certainly be to take, as a sample naturalistic ethical theory, some highly plausible theory rather than act consequentialism. After all, Adams considers the critical stance to underwrite a critique of naturalistic ethical conceptions generally. How about this bit of ethical theory: T\*="It is wicked to torture children"?

Suppose that empirical evidence becomes available which, by the ethical naturalist's standards, confirms T\*. Would it then be rationally required that one accept T\*? Would it remain an open question whether or not T\* is correct?

I think that the ethical naturalist is obliged to deny that there is any sense in which it would be, by her standards, an open question whether or not T\* is correct, except the sense which corresponds to a critical-stance-on-the-cheap: the sense in which any conclusion not analytically entailed by empirical evidence is an "open question." On the other hand, I do not think that the naturalist should be embarrassed by this conclusion.

Let's consider what sort of empirical evidence would, by the naturalist's standards, confirm T\*. Now, of course, I have chosen T\* so that its plausibility makes it a candidate for quasi-analyticity. In general, *a posteriori* quasi-analytic claims receive most of their empirical support through their inferential connections to other well confirmed claims, so that it is often hard to say what particular bits of empirical evidence most straightforwardly support them. In the case of T\*, however, I think that we can say what sort of empirical evidence might properly be thought to be especially strongly confirmatory.

In general, empirical evidence counts significantly towards the confirmation of a theory just in so far as it tends to rule out one or more of the theoretically plausible (that is: projectible) alternatives to it (Boyd 1985, 1985a, 1991). That's why it is often hard to say what evidence specifically supports a quasi-analytic claim: often there is no theoretically plausible alternative. In the case of T\*, however, I think that we can identify two alternatives whose plausibility (by an ethical naturalist's standards) is sufficiently high to permit the identification of specific potentially confirmatory sorts of empirical evidence.

Here's what I have in mind. The judgment that torturing children is wicked presupposes that there is such a phenomenon as wickedness, and it asserts that torturing children has such properties as are required to fall under the category *wickedness*. It is, I think, (just barely) credible that one or both

of these claims is mistaken. Each would be addressed by the empirical confirmation, by naturalist standards, of T\*.

Consider what would be required for such an empirical confirmation. First, it would be established, by naturalist standards, that there was some natural phenomenon to which the term “wicked” refers. Let’s use Adams’ version of the naturalistic conception of reference to see what this would mean. The term “wicked” plays an inferential role in ethical discourse which presupposes a certain nonarbitrariness or unity to the things we classify as wicked. We have systematic expectations about the ways in which wicked behaviors are related to human goods and harms, to features of moral character, to patterns of moral development, and to each other. The confirmation, by naturalist standards, of the claim that “wicked” refers *at all* would require confirmation of the proposition that there is a family of natural properties such that the things that poses them answer to this role and such that our judgments about which actions (policies, personalities, etc.) are wicked pretty reliably track this family of properties. Were this proposition confirmed, then by the naturalist’s standards it would be established that that family of properties defines the real category to which the term “wicked” refers.

What about the remote epistemic possibility that torturing children does not satisfy the natural definition of “wicked”? How could this happen. Well, it *could* turn out that the empirical facts are very much different from what we think and that the situation of the practice of torturing children within the causal order is relevantly different from that of really wicked things. Perhaps some form of child torture is crucial to the moral development of both children and their torturers and perhaps there are higher artistic and intellectual pleasures which are only available to those who have participated (as both torturer and torturee) in child torture. Perhaps only torture can protect children against some utterly (physically, morally and psychologically) debilitating condition in later life....

One shouldn’t stay up nights worrying about such possibilities; that’s why T\* is quasi-analytic. But, at least from a naturalist perspective, these (epistemic) possibilities represent the only even remotely intelligible ways in which T\* could turn out to be false (assuming that there is such a phenomenon as wickedness). Of course the hypotheses that these possibilities obtain are empirically testable hypotheses about the natural world, so when we assume that T\* has been confirmed by the naturalist’s standards we must assume that they have been disconfirmed by empirical evidence.

OK, suppose that, and suppose that it has been confirmed—again by the naturalist’s standards—that “wicked” refers to a real property. Would it then be rationally required that one accept T\*? Well, from a naturalist’s perspective (more on the epistemology of ethical naturalism next) the answer must be “yes.” Is this a problem for the naturalist? Is there some way she has missed

in which the question of T\* is an open question? Is it part of our concept of moral epistemology that T\* must—under the conditions just described—be an open question (not on-the-cheap)?

It seems to me that, if our ordinary conception of the epistemology of morals dictates an answer to these questions, it is the opposite of the one required by Adams' critical stance. After all, I have chosen as T\* a claim which Adams apparently considers better established than lots of fundamental and well established physical theories. I take it to be a virtue of naturalistic ethical conceptions (no doubt shared by some important non-naturalistic conceptions) that it explains why T\* is quasi-analytic and not the subject of an open question.

I do not deny that there are coherent and historically important conceptions of the epistemology of morals according to which it would be an open question whether or not T\* is right, given the empirical evidence envisioned. As far as I can see, all of these conceptions are theistic: they involve something like (relatively implausible versions of) the divine command conception. If one of these conceptions is right, then Adams and the rest of us are mistaken to treat T\* as largely immune from refutation. Still, the existence of such conceptions does not establish what Adams needs: that the critical stance is part of our current conception which an adequate metaethical perspective must rationalize. If anything, the opposite is the case.

1.3. *Naturalism and the Epistemology of Morals.* I have been addressing the question of the critical stance as though it is part of the naturalist realist conception in metaethics that, if a naturalist conception of the nature of some moral property (goodness, wickedness, etc.) is confirmed by empirical evidence, the confirming evidence consists of evidence which indicates that the nature in question plays (at least quite well, and better than any other candidate *natural* phenomenon) the inferential role associated with the relevant ethical term. This seems to be how Adams understands naturalist realism, so I have not begged any questions against him in adopting this approach.

Nevertheless, I think that the naturalist realist conception of the epistemology of morals is more complicated, and that, when the complication is made explicit, two interesting points about the metaphysics, epistemology and semantics of morals become clear. First, Adams is right that there is a certain sort of critical stance which is appropriate when we examine substantive moral issues in the context of the metaphysics of morals. I think that in fact neither naturalistic moral realism nor theistic moral realism has any difficulty in accounting for this sort of critical stance, but I think also that a tacit appreciation of it may (mistakenly) lend credibility to Adams' critique of naturalism.

The second point is the one advertized earlier: that Adams' discussion of the critical stance raises crucial questions about the metaphysics of semantic theory. I'll explore the second of these points in Part 2. For the present, I want to explore the first one.

Here's what I have in mind. We lack a fully developed theory of "natures" adequate to address every subtle question regarding our identification of them, but one thing seems clear: it is at least an important constraint on the identification of the nature of a thing, property, kind, or whatever, that we should, *prima facie*, look for the most fundamental, or metaphysically deepest, candidate which suitably fulfills the inferential role associated with the relevant term (as understood in the light of the epistemic access and achievement explanation conditions). That's why, for example, we are inclined to take its atomic number as the nature of an element rather than simply the family of causal powers summarized by its position in the periodic table. The former is more fundamental than the latter, because it explains those causal powers in a unified way.

For this reason, I suggest, ethical naturalism is, to a good first approximation, an *atheist* (or perhaps *atheist or deist*) position. Suppose, to choose an example which is more plausible by Adams' standards and by mine, that empirical evidence establishes it beyond scientifically reasonable doubt that a homeostatically united cluster of human-flourishing-related properties fulfills quite well the role associated with the moral uses of the word "good" in just the way some defender of indirect consequentialism suggests, and that it does so much better than any other natural phenomenon. Is it a plausible ethical naturalist's position that this evidence, *by itself*, would settle the question of the nature of the good in favor of naturalistic homeostatic property cluster consequentialism?

It seems to me that the answer is plainly "no." Suppose that there is a Deity such that Her command is that we love and care about each other and such that the homeostatic property cluster in question—and the natural phenomena corresponding in a reference-like way to other moral terms—are reflections of Her implementation, in the created world, of human well being. Surely then, the nature of the (moral) good is something metaphysically related to Her (perhaps it *is* She). The *natural* phenomenon most closely corresponding to the role associated with the term "good" bears to the real nature of the good something like the relation which the causal powers indicated by carbon's place in the periodic table bears to the property of having atomic number 6.

So, in order to defend ethical naturalism, the naturalist must deny that any such supernatural nature is available to provide the deeper "fit" to the inferential role associated with the word "good." The most straightforward way to do this—the one associated systematically with the term "naturalism"—is to

deny that there are any supernatural phenomena at all. Of course, other options are possible, so long as they entail that such supernatural phenomena as there are do not bear some special (role fulfillment explaining) relation to moral practice.

In any event, the sort of evidence about the relation between the use of "good" and natural phenomena which would lead a philosopher *who is already a naturalist* to think that the nature of the good is the particular natural phenomenon she identifies would not, *even by her own standards*, be sufficient *by itself* to establish her naturalistic hypothesis about the referent of "good". Additional evidence would be required in order to rule out a more fundamental supernatural explanation for the relation in question. Of course the naturalist moral realist will insist that such evidence is available. Perhaps it involves the increasing success with which natural phenomena are explained in a purely materialist way. Perhaps the Darwinian revolution in biology is crucial in undermining natural theology.

Whatever her position, it will certainly not be the case that she thinks the way in which some natural phenomenon fits the role associated with "good" settles the issue of theism. So, on her view, it will be, given the evidence of such a fit, an open question what the nature of the good is. That evidence will to some extent constrain theories of the good, but it will not rule out a theistic conception.

Evidently, neither the theist nor the atheist—and thus neither the naturalistic moral realist nor the theistic moral realist—will have any trouble explaining this sort of critical stance with respect to the epistemology of morals. Nevertheless, I think that this sort of case may mislead the theistic moral realist with respect to issues about the critical stance. Adams suggests that she should ask herself, "Would I be rationally persuaded of X, given the empirical evidence which would convince an ethical naturalist moral realist of X," for X some naturalistic moral theory.

If she interprets this question as "Would I be rationally persuaded of X given the empirical evidence which would persuade *someone already committed to some version of ethical naturalism* of X?", the answer will, of course, be "no," and she may interpret this result as indicative of some special critical stance especially appropriate to ethics.

Instead, the question she should consider is this one: "Would I be rationally persuaded of X given the empirical evidence which convinces philosophical naturalist moral realists to reject the supernatural, *together with* the evidence which would persuade someone committed to ethical naturalism of X?" The answer to that question might also be "no," but that would indicate, not something peculiar about the epistemology of morals, but, if anything, something about the epistemology of religion. It is something which it is not the special task of any metaethical theory to explain.



1.4. *Postscript on the Metaphysics of Criticism.* I want to turn presently to the deep questions about reference which I believe are raised by Adams' discussion of the critical stance. Before that, I want to make one more cautionary remark about the practice of relying on epistemic judgments or intuitions in exploring issues of moral epistemology. Adams suggests in Chapter Two that the naturalist moral realist might try to explain the epistemic judgments which Adams takes to support the critical stance as reflections of the general uncertainty of belief formation in the empirical sciences. With the qualification just articulated about the epistemology of religion, I think that this is basically right, and that considerations about the uncertainty of knowledge of complex psychological, social, political and economic factors—especially in the light of the influence of social ideology on moral and social inquiry—are sufficient to justify whatever sort of critical stance one should have regarding matters moral.

Adams suggests that his own theistic account of the good could provide a metaphysical explanation (and justification) for our epistemic uncertainty regarding moral inquiry. If God's metaphysical nature is such that She is, in important respects, unknowable, then—on Adams' theory—some of that unknowability carries over into the realm of moral inquiry. If this approach is right then there is a metaphysical underpinning to our epistemic caution in moral inquiry.

I want to make two points about the metaphysics of this sort of caution. In the first place, note that the naturalist's explanation for our warranted caution in such matters is itself metaphysical: it is a matter of how we, as knowing subjects, are situated in a complex web of physical, psychological and social causes. If facts about the systematic relations of our epistemic practices to causal structures are not metaphysical facts, then it's not a metaphysical fact that water is H<sub>2</sub>O.

So, if you think that epistemic facts must rest on metaphysical foundations—and you should—then you need to recognize that what we have here are two competing metaphysical explanations. Of course, the naturalist's explanation is less elevated, but that's the fate of naturalistic metaphysics.

The other point is about the "resolving power" of our epistemic judgments or intuitions. I suggest that there is no reason to suppose that such judgments, being artifacts of our particular social and intellectual upbringing, are up to the task of providing evidence for or against any particular account of the metaphysics of moral doubt. The factors which influence our judgments about the epistemology or morals are so diverse—ranging from theological and political convictions, through various degrees of childhood exposure to culture relativism, to philosophical opinions on the epistemology of other sorts of inquiry—that almost any plausible philosophical theory of the semantics and metaphysics of moral discourse can reasonably explain them.

Instead of relying on those intuitions, I believe, we should do the relevant metaphysics. If you have (largely) independent reasons to accept theism and a theistic conceptions of ethics, then you have reason to accept the deeper metaphysical explanation and justification for moral doubt. If you have (largely) independent reasons to accept philosophical naturalism, then you have reason to accept the less elevated metaphysical explanation.

## **2. The Semantics of “Good” and of Other Natural Kind Terms.**

2.0. *Adams on Inferring “Is” from “Ought.”* Recall that when Adams considers the (very, very, very) counterfactual situation in which empirical evidence, as a philosophical naturalist would assess it, supports act consequentialism, he concludes that he would not feel rationally compelled to accept that theory, and he explores various options for the naturalistic moral realist in accounting for this judgment. We have just examined the response which, I argue, the naturalistic moral realist should give.

That response does not deploy one of the semantic options which Adams considers on behalf of his opponent. According to Adams, faced with evidence which she considers to have confirmed some naturalistic moral theory, and with someone like Adams who seems justified in failing to agree, the naturalist moral realist might invoke the notion that the term “good” and, presumably, lots of other moral terms *partially denote*. In so far as I understand it, the suggestion is that the naturalist might respond to her failure to convince Adams by positing *two different moralities*, one appropriate to her naturalistic account of the nature of the good, the other to Adams’ (equally rational) alternative account.

I have already indicated why I don’t think this is the appropriate response. The naturalist would be justified by her own standards in taking her theory of the nature of the good to be confirmed only if it were the only plausible *naturalistic* theory compatible with the empirical data. Thus she would have to conclude that the theory Adams prefers could only be underwritten by a supernaturalistic metaphysics, which she takes herself to have (largely) independent reasons for rejecting. So, she should conclude (with the tentativeness and caution appropriate to empirical inquiry about complex politically controversial topics) that there is only one morality—the one appropriate to the nature she has identified as the (ethical) good.

Of course, if she considers hypothetically the possibility that Adams’ theory (or some other theistic theory not fatally compromised by the empirical data) is correct, then under that assumption she could hold that *something like* partial denotation is involved. “good” refers to some theistically defined phenomenon but it bears a *reference-like* relation to the natural phenomenon in question, just as “carbon” bears a reference-like relation to the family of causal powers indicated by a particular position in the periodic table. But this

is not a two-moralities theory; it does not save the day for her conception of the nature of the good. Instead it demotes it to the status of a metaphysically unrevealing (albeit morally informative) approximation.

Nevertheless, I think that what motivate Adams' consideration of this alternative for the naturalistic moral realist are two real insights about the epistemology and semantics of moral discourse. The first of these, I have mentioned favorably earlier. Part of what Adams is insisting on in the chapter in question is that moral judgments can "trump" scientific judgments. In particular, if we are faced with evidence which seems to establish that the term "good" refers to some natural phenomenon, P, the moral implausibility of P's being the good can in principle undermine or override such evidence. Moral inquiry has a certain level of autonomy with respect to other areas of inquiry. If one thinks of this autonomy as absolute, then the appeal to partial denotation might be one way to save some sort of moral realism without adopting Adams' solution.

Of course, as we have seen, the naturalist moral realist should not think that moral inquiry is absolutely autonomous, so there is no reason why she should invoke partial denotation under the circumstances in question. I want to emphasize, however, that the non-reductionist naturalist moral realist will agree with Adams that there is no scientific discipline whose findings have some sort of absolute epistemic priority over our moral and ethical judgments. For the naturalist, after all, ethical inquiry is simply one domain of empirical inquiry, methodologically (*and ontologically*) continuous with psychology, social theory, economics, history, etc. For the non-reductionist naturalist, there is no special priority which any one of these disciplines has over the others. In fact, given the special vulnerability of inquiry in all of these domains to the influence of pernicious social ideology, it would be unwise to assign any special methodological role to (institutionally certified) expertise in any of these disciplines including ethics itself.

Naturalist moral realism is thus not (or not necessarily) a scientific position. It does not entail that there is some scientific algorithm for deciding ethical questions. Adams is right to discern that he and I differ about the extent to which moral reasoning and scientific theorizing should resemble each other. The difference stems, not from our differences about the ultimate metaphysics of morals, but, I believe, from different estimates about the sorts of problems we face in trying to care for each other. I'm inclined to believe that these problems stem largely from unfortunate features of social, economic and political structures, so I'm inclined to think that moral theory is in large measure a branch of political economy, *properly done* (which is, very definitely, not to say that current institutionalized "social science" research has a significant contribution to make to moral theory). Someone with a different estimate of the moral problems we face could be equally a naturalist

moral realist but place much less emphasis on the similarities between moral inquiry and scientific theorizing.

*2.1. Normativity and Objective Reference for Moral Terms.* I turn now to the second important point raised by Adams' discussion of naturalist moral realism and the critical stance. One thing which his discussion of the hypothetical situation in which evidence seems to confirm act consequentialism indicates is that judgments about the semantics of ethical terms are, in part, ethical judgments. Adams' confidence that he would not feel rationally compelled to accept act consequentialism, given the scientific evidence he envisions, arises from his recognition that it is an absurd moral theory. The fact that the ethical naturalist can concur in that judgment eliminates one threat to her position, but it does not address issues for the naturalist arising from the fact that semantic issues regarding ethical terms are themselves partly ethical issues.

Of course, as Adams points out, that's not surprising, since questions about the reference of chemical terms are partly questions in chemistry. Nor does it pose a problem of circularity for the moral realist since, as Adams notes, she is not trying to introduce moral terms by stipulative definitions. Nevertheless, the fact that questions about the reference of moral terms are so closely related to normative ethical questions raises an important challenge to the moral realist who is also a philosophical naturalist.

Here's why. As Adams points out, many philosophers accept an internalist conception of moral judgments according to which accepting the judgment that something is (morally speaking) good must, by itself, provide some reason for choosing or preferring it. As I indicated, it is on this point where Adams comes closest to holding that the reference-determining role associated with a referring expression is specifiable analytically, or by conceptual analysis. He appears to treat it as analytic that, if goodness has a nature, then that nature underwrites this universal reason-providing role for judgements about which things are good. Pretty obviously, the theistic conception of the good Adams offers is a reasonable candidate for just that role.

Famously, it is hard to defend internalism and naturalism simultaneously, while affirming an independently plausible substantive moral theory. The standard naturalist alternative is to identify some candidates as the referents of "good" and of other moral terms and then to argue that the hypothesis that these are the referents of the terms in question explains—by reference to facts about typical human preference structures—why we have the mistaken impression that moral judgments are internally related to reasons for action.

Consider now the impact on that strategy of the insight of Adams' that judgments about the referents of moral terms are themselves partly moral judgments. In identifying the referent of, say, the term "good" the naturalist

will have had to rely on just the sort of moral judgments which, according to her internalist opponent, have inferential properties not explicable within a naturalistic framework. If the identification of the nature of the good proceeds by methods which a naturalist cannot rationalize, then the subsequent eliminative rationalization of internalist intuitions may be unconvincing.

It is important to see that concerns of this sort do not arise only in the (bizarrely) hypothetical situation in which empirical evidence might seem to support so implausible a theory as act consequentialism. On any plausible naturalistic conception of reference, the relation which use of the term "good," in ethical contexts, bears to the family of phenomena which are favorably evaluated by prevailing ethical standards within the relevant society must be *something like* the reference relation. This is why ethical relativism is an initially plausible semantic conception. Presumably, part of the moral realist's justification for rejecting the relativist claim that the relation in question *is* the reference relation is the *moral* conviction that prevailing moral conceptions are often in error. Apparently moral judgements really matter in theorizing about the semantics of moral terms, even when bizarre hypothetical possibilities are set aside.

I am not suggesting that these considerations make the naturalist's approach viciously circular. For reasons which both Adams and I accept, one can see that, if the position the naturalist eventually articulated encompassed a sufficiently convincing rebuttal to internalist intuitions, a successful defense of naturalist moral realism might be possible. Still, it would be nice for the naturalist to have at least a preliminary account of the nature of reference determination which points toward a non-internalist account of referential judgments for moral terms.

There are other reasons why providing such an account may be seen to be a pressing problem for the naturalist moral realist. In the first place—whether or not we accept the possibility of the *human* amoralist—it would seem that it should be possible for non-human linguists, without anything like moral commitments, to investigate the semantics of human languages, including their moral components. Such investigators would plainly not be making moral judgments in any ordinary sense, so the naturalist faces the problem of indicating how they could determine the referent of, for example, moral uses of "good."

There is also the interesting question of how non-empiricist theories of the semantics of natural kind terms (and other terms like "good" if it happens to be, for example, a supernatural kind term) should be understood. In exploring the critical stance, Adams envisions a situation in which the evidence which a naturalist moral theorist takes to confirm a general moral theory consists primarily of evidence about the ways in which various natural phenomena are related to the uses of moral terms. If this picture is basically right,

then a naturalistic investigation of the nature of the good would be, in a certain sense, a matter of investigating the uses of "good." Of course, as we have just seen, such an investigation would have to take place in a context determined by, among other things, moral judgments. Even so, the investigative strategy Adams envisions would assign a significant methodological role to data about language-world relations.

This is an interesting, but odd, idea. On the one hand, Adams is right to suggest that appeals to semantic facts about "good" and other moral terms do seem to play a significant role in the defense of particular naturalist moral theories, and not just because one needs something like a naturalistic theory of reference for such terms in order to avoid the trap of thinking that they must have analytic definitions.

On the other hand, defenders of realist naturalism about real rather than nominal "natures" in the scientific domain have not generally thought of their proposals as entailing that scientists should engage in anything like *language-centered* investigations. Chemists found out that the nature of water is the structure described by the formula " $H_2O$ " entirely without help from metaphysicians and philosophers of language, by investigating water not "water." I think that part of the plausibility of Adams' critique of naturalistic moral realism arises from the peculiarity of the idea that we could somehow do ethical theory by doing linguistics or philosophy of language.

Thus the plausibility of naturalistic moral realism would be greatly enhanced if there were a plausible naturalistic account of the epistemology of morals which simultaneously (1) provides a non-internalist account of the role of moral judgments in establishing semantic facts about moral terms and (2) does so without assigning an implausibly great methodological role to purely semantic inquiries about moral terms.

We may put the challenge to the naturalistic moral realist this way. Judgements about the semantic properties of moral terms depend on normative moral judgments. So if, as naturalist moral realists believe, there is an objective fact of the matter about, e.g., the reference of the term "good," closely connected to objective facts about the nature of the good, then it would appear that there must be such a thing as *objective normativity* of the sort which internalists affirm. But plausible naturalistic versions of moral realism—unlike theistic versions for example—are not compatible with internalism.

*2.2. The Problem Generalized: Normativity and the Semantics of Scientific Terms.* In order to address this challenge, we need to see something about its scope. Adams situates the problem of the role of normative judgments in the semantic theory of moral discourse in the context of naturalistic vs. non-naturalistic conceptions of the good. I'll suggest, by contrast, that the problem of

the role of normative considerations in semantic theorizing extends to theories of the semantics of natural kinds in the sciences as well. As it happens however, this problem has a naturalistically acceptable solution, and that solution works for the special case of the semantics of moral discourse.

Here's why the problem generalizes to the scientific case. In the case of any natural kind term, *t*, which refers to some natural kind, *k*, there will always be other sorts of things (kinds of a sort) to which *t* also bears a reference-like relation. At any given time, the category of things which satisfy the prevailing standards for the application of the term *t* will be one such category. The correspondence between *t* and this category need not be unimportant intellectually. For example, suppose that it is correct to say, as most taxonomists now do, that birds (the traditional class *Aves*) are part of the dinosaur taxon *Maniraptora*. Suppose that you are reading a work by an eminent paleontologist on "dinosaurs of the American West" which was written before this view became credible. If you want to learn about paleontology from this book, you will be well advised to take seriously the quasi-referential connection between the term "dinosaur" and (what is not its referent) the paraphyletic taxon that consists of *Dinosauria* with *Aves* omitted.

In the case of scientific terms which play a central role in disciplines significantly influenced by social ideology, there will often be semi-referents defined in terms of ideological role. Finally, in almost every sort of case there will be candidate referents which don't make the grade but which certainly bear a reference-like relation to the term in question. For example, for any element with one principal isotope, its chemical name certainly bears a reference-like relation to that isotope. Once you're on a roll, you can generate examples like this almost without limit.

In none of these cases does the verdict in favor of the right referent follow from simple statistical facts about verbal behavior, like facts about which category affirmations of kind membership most closely tracks; in this regard the situation is like that regarding the referent of "good." That is no surprise: in each case when we inquire about reference we understand ourselves to be inquiring about "natures," and the whole point about natures is that they are supposed to reflect deep or nonsuperficial features of reality.

What is important for our purposes is that the similarities run deeper. In deciding what the referent is of a natural kind term, we are seeking to identify, from among the candidate categories, the one which *best fits* the explanatory role associated with natural kind terms: explaining the inductive, explanatory and practical achievements of the associated discourse. In a perfectly good sense of the term, we are making *normative* judgments here. Similarly, referential hypotheses have normative implications, of a sort. Since such hypotheses are (components of) explanations of the ways in which language use contributes to the *achievements* characteristic of the relevant discourse

practices, they imply hypothetical normative judgments about how to bring about such achievements. Hypotheses about the natures of chemical kinds, for example, imply normative judgments of the form, "In order to obtain inductive, explanatory and practical achievements of the sort characteristic of chemical practice, one must classify reagents according to a scheme which is (relevantly like) the following...."

I'll defend a conception of the metaphysics, epistemology and semantics of natural kinds and of natural kind terms according to which the only normative judgments involved in the evaluation of semantic claims about natural kind terms are *epistemic* judgments about the cogency of competing explanations for achievements within particular domains of practice, *and* according to which the only normative judgments which are implied by such semantic claims are hypothetical judgments about how to bring about such achievements. Thus, for example, the normative judgements involved in determining that the nature of water is  $H_2O$  and the nature of carbon is having atomic number six are purely epistemic judgments about the merits of competing explanations of the ways in which language-world relations contribute to inductive and explanatory successes in chemistry and related disciplines. Similarly the only normative implications of such semantic hypotheses involve the sort of hypothetical normativity amenable to naturalistic interpretation (in imperitival form, roughly: "Classify this way if you want to discover true generalizations about how chemical reactions work.")

Of course, what I am going to insist is that the objectivity of these sorts of normativity is naturalistically explicable *and* that it is open to the naturalist moral realist to argue that only these sorts of normativity are implicated in investigations of the semantics of moral terms and of the metaphysics of moral. In particular, it is open to the naturalist moral realist to maintain that the normativity of assertions about the metaphysics of morals is just the same sort of hypothetical normativity which attaches to hypotheses about the metaphysics of chemical kinds ("Classify this way if you want to discover true generalizations about how chemical reactions work; classify that way if you want to figure out how we can effectively care about each other's well being"). To argue for this conclusion I'll need to explore the metaphysics and semantics of natural kind terms.

*2.3. Kinds as Constructions and Referential Hypotheses as Achievement Explanations.* The special role of judgments about goodness of fit in assessing referential hypotheses illustrates a problem which has plagued attempts to formulate so called "pure" causal or naturalistic conceptions of reference, even for scientific terms. If one thinks of kinds as phenomena "out there" in the world independently of our concepts and practices, and if one then holds that reference is a matter of a certain sort of causal relation between terms in use



and those independently existing kinds, then there will be lots of different kinds causally associated with the use of any given term, and it is hard to say what distinguishes its referent from the others.

The solution to this problem derives from two important insights into the metaphysics of natural kinds and the semantics of natural kind terms. The first is that the notion of a natural kind is intimately associated with the notion of a causally sustained regularity. *Natural* kinds are just those we need to refer to in order to formulate *projectible* hypotheses, and these in turn are those which are candidates for being *causal laws* or (better) *causal generalizations*: laws or generalizations reliably sustained by causal processes or mechanisms. The role of reference to natural kinds is to achieve an *accommodation* between our conceptual and inferential practices and relevant causal structures. Call this the *accommodation thesis* (Boyd 1991, 1992, 1993, 1999a, 1999c, 2001b).

The second important insight (call it the *discipline relativity thesis*) is that the “naturalness” of a natural kind is relative to the discipline or disciplines within which reference to it serves the function identified by the accommodation thesis. It is this insight which philosophers express, for example, when they say that *pain* is a natural kind “from the perspective of psychology” but probably not (owing to multiple realizability) “from the point of view of fundamental physics.”

What the accommodation thesis and the discipline relativity thesis suggest is that the philosophical theory of natural kinds is a theory about how human practices (or the practices of other representational systems) are able to accomplish the achievements associated with particular disciplinary regimes. Natural kinds, according to this *achievement explanation thesis*, are in an important sense artifacts or social constructions of the human practices involved in achieving accommodation. Locke maintained that while Nature makes things similar and different, kinds are “the workmanship of men.” If the achievement explanation thesis is right, then (gender bias aside) he was right: A natural kind *just is* the implementation, in language and in conceptual, experimental and inferential practice, of a (component of) a way of satisfying what we may call the *accommodation demands* of some matrix of human practices.

Locke erred not, as some interpretations of naturalism about kinds would have it, in seeing kinds as matters of our workmanship, but in his conventionalism about the character of that workmanship. Locke said that “...each abstract idea, with a name to it, makes a distinct Species.” His conception was that kinds are established by a sort of *unicameral* linguistic legislation: people get to establish kind definitions by whatever conventions (nominal essences) for the use of general terms they choose to adopt.

According to the accommodation thesis, natural kinds are instead products of *bicameral* legislation in which the (causal structure of the) world plays a heavy legislative role. To a good first approximation, a natural kind is nothing over and above a natural kind term together with its use in the satisfaction of the accommodation demands of a disciplinary matrix. Or, better yet, the *establishment* of a natural kind (remember that natural kinds are legislative achievements—that is, artifacts) *consists solely in* the deployment of a natural kind term in satisfying the accommodation demands of a disciplinary matrix.

Natural kinds are features, not of the world outside our practice, but of the ways in which that practice engages with the rest of the world. Biological taxonomists sometimes speak of the “erection” of biological taxa, treating such taxa as, in a sense, human constructions. They are right—and the same thing is true of natural kinds in general.

All of this is pretty metaphorical, so here’s how to make it more precise. I’ll use the term *disciplinary matrix* to describe the body of aims, methods, theories, concepts, etc. associated with a domain of everyday, or scientific, or moral inquiry and practice. To a good first approximation, ignoring issues like partial denotation about which more later, here’s the theory of kinds, natures and reference appropriate to the conception I have been articulating.

Let  $M$  be a disciplinary matrix and let  $t_1, \dots, t_n$  be the natural kind terms deployed within the discourse central to the inductive/explanatory successes of  $M$ . Then the families  $F_1, \dots, F_n$  of properties provide explanatory definitions of the kinds referred to by  $t_1, \dots, t_n$ , and determine their extensions, just in case:

1. (Epistemic access condition) There is a systematic, causally sustained, tendency—established by the causal relations between practices in  $M$  and causal structures in the world—for what is predicated of  $t_i$  within the practice of  $M$  to be approximately true of things which satisfy  $F_i$ ,  $i=1, \dots, n$ . In particular, there is a systematic tendency for things of which  $t_i$  is predicated to have (some or most of) the properties in  $F_i$ .<sup>6</sup>

2. (Accommodation condition) This fact, together with the causal powers of things satisfying these explanatory definitions, causally explains how the use of  $t_1, \dots, t_n$  in  $M$  contributes to accommodation of the inferential practices of  $M$  to relevant causal structures. It explains whatever tendency there is for participants in  $M$  to identify causally sustained generalizations, to obtain correct explanations, or to obtain successful solutions to practical problems.

---

<sup>6</sup> Think of predicating  $t_i$  of something some expression,  $a$ , as predicating “...has  $a$  as a member” of  $t_i$ .

It is this conception of reference and of (the natures of) kinds which I propose to rely on in discussing the role of normative judgments in assessing referential hypotheses.

*2.4. Roles, Disciplinary Matrices and the Real Form of (Referential) Semantic Hypotheses.* I want first to situate the proposed account of reference and of natures by contrasting it with Adams' account. As Adams suggests, a term, *t*, within the context of a particular domain of inquiry or practice, refers to a kind (or whatever) whose nature suits it to a particular role, *but*, according to the account I am proposing, the role in question is determined, not by the concept of *t* held by practitioners, but instead by the way their use of *t* contributes to their actual achievements. There are thus two different ways in which the practitioners' concept of *t* could be mistaken, and thus two reasons to deny that reference fixing roles can be identified by conceptual analysis. First, practitioners within a practical or theoretical discipline can be, and often are, wrong about what it is that they accomplish. Second, they may be wrong about the ways in which their conceptual and methodological practices contribute to the accomplishments of which they are capable.

It follows that the analysis of the natures to which natural kind terms refer is not a matter of "conceptual analysis" as philosophers ordinarily understand it, even though techniques of conceptual analysis may be (epistemically fallible) components in our methodology for investigating natures.

Consider next the content of statements of the form "*t* refers to the kind with nature *N*." Note first that statements of this sort express a definite proposition only when the context of utterance specifies sufficiently clearly a disciplinary matrix with respect to which the question of the reference of *t* is posed. When this context dependent specification is taken into account we can see that the propositions that are expressed by such statements are species of the genus of propositions of the form "Epistemic access to things with *N* reflected in the use of *t* *within practical or theoretical discipline D* explains aspects *A* of the achievements of practitioners within *D*." Cases of what I have been calling quasi-reference resemble cases of reference precisely in that, when *t* enjoys a quasi-referential connection to the kind with nature *N*, *t* and *N* satisfy the formula just enunciated with respect to some specifications of *D* and *A*.

The distinction between quasi-reference and reference is not sharp, but roughly situations in which the formula in question holds for *t*, *N*, *D* and *A*, but *t* does not refer in *D* to the kind with nature *N*, will be ones in which one or more things like the following happen:

1. The same relationship holds between *t*, *N'* (*N' ≠ N*), *D* and *A'* (where *A=A'* or where *A'* includes all or most of the important achievements in *A* together with other achievements of *D*) so that the *t-N'* relationship explains

more of the success of practice within D than does the t-N relationship (in which case the referent of t in D might be the kind with nature N').

2. There are N' ( $\neq$ N) and A' ( $\neq$ A) such that the relationship in question also holds between t, N', D and A' where the importance of each of these relationships is such that, in D, t *partially denotes* (in the sense of Field 1973) both the kind with nature N and the kind with nature N'.

3. For every N\* such that, for some A\*, t bears the relation to N\*, D and A\*, the depth of the explanation provided for A\* falls short of that ordinarily associated with attributions of reference, so that it is preferable to say that t, in D, fails to refer.

A clear example of the first sort of case can be generated by fixing some particular time in the history of an epistemically successful and rapidly changing scientific discipline and considering the relation between some natural kind term in that discipline and the family of things which *at that very time* would be classified under that term by the standards then prevailing. Relations of this sort figure in the explanation of disciplinary success, because there is a systematic tendency in successful sciences for classificatory judgments at any given time to be a pretty good indication of causal factors relevant to the phenomena under study. Even so, a *more complete* explanation for the same success will ordinarily make reference to the correspondence between classificatory terms and somewhat different families of properties—ones to which the families picked out by instantaneous time-slices of practice are approximations. These latter families will be the natures to which the relevant terms refer. The correspondence between the terms in question and those families explains more aspects of the relevant epistemic success in greater detail.

To see how cases of this sort arise (and to see how issues about reference depend on choice of relevant discipline), consider the term “acid” as it was used in Renaissance chemistry. It is plausible that for a long time indeed the prevailing standards for identifying reagents as acids were given by a short list of properties which were both diagnostic and practically relevant: acrid smell, sour taste (if diluted), corrosiveness, and a few others.

In an ordinary philosophical context it would be right to say that the term “acid” refers—and *referred then*—to the kind whose nature is given by a different property, say having pH less than seven in aqueous solution.<sup>7</sup> This is so because in such contexts the question of the referent of “acid” is the question of what nature there is, such that its reference-like connection to the term

---

<sup>7</sup> Having pH less than 7 is not really a very good definition for acidity. Consulting advanced chemistry texts shows that there is a variety of more sophisticated proposals in the literature. The point developed here can be made with respect to any of them.

“acid” best explains the epistemic reliability of chemical practice *over the long haul*. This, in turn, is so because, in the standard philosophical context in which we inquire about the natures of chemical kinds, we have it in mind that the chemical practice we are interested in is historically extended, running (at least) from Renaissance chemistry through current chemical practice and into the foreseeable future. With respect to the reliability of that practice, the association between the term “acid” and low pH is explanatorily crucial; it is even important that all of the reagents that early chemists called “acids” have this property and that they exhibited the features those chemists took to be diagnostic of acidity because they have it.

On the other hand, if we raise the question of the reference of “acid,” not in the long (and continuing) history of chemistry, but in Renaissance chemistry understood as an historically delimited body of practice, then it may well be correct to say that the kind to which the term “acid” referred was that defined by the then standard diagnostic properties. If we restrict our attention to periods prior to the emergence of modern chemistry, there are many fewer epistemic successes to explain, and the invocation of pH may be explanatorily irrelevant. Such is the context dependence of the truth conditions for statements of the form “t refers to k.”

Cases of the second sort are routine in science; the standard example involves the term “element” and the terms for various elements before the discovery of what (after denotational refinement) we call “isotopes.” The third sort of case is probably illustrated by the term “caloric,” which most philosophers treat as having been non-referring despite its reference-like connection (within physics) to conductive heat transfer.

By now, you may wonder what all this has to do with normativity and, ultimately, with the nature of the good. Here’s the idea. I suggest that what examining the accommodation and epistemic access conditions, and recognizing the form of the propositions expressed by attributions of reference to kinds defined by natures, allows us to see is that the only sort of objective normativity there needs to be in order to underwrite objective judgments about reference, kinds, and natures is objective *epistemic* normativity.

As the accommodation and epistemic access conditions indicate, what’s at issue in questions of reference to kinds defined by “natures” is the establishment of correct and suitably complete explanations for the epistemic reliability of various human practices. Issues about the correctness and completeness of explanations are issues addressed by epistemic norms. In particular, when we distinguish between cases of quasi-reference and cases of genuine reference, the standards we are applying are, as the previous discussion indicates, matters of the accuracy and completeness of explanations *and of* the appropriateness of those explanations to the particular aspects of disciplinary epistemic reliability indicated by the context in which questions of reference arise. The

impression that judgments of “depth” or “importance” are involved which go beyond epistemic questions about the adequacy of explanations derives from two sources.

One is simply that these are terms we use to describe explanations which are more complete, in ways relevant to the questions being addressed, than their rivals, so that when we reject a referential hypothesis that explains less about the reliability of practice than does some rival we *are* making a judgment about explanatory depth or importance, but this is an epistemic matter.

The second factor concerns the typical concerns (pun!) of philosophers. Ordinarily, and for good reasons, we are concerned to examine the metaphysics and epistemology of successful, *long term*, theoretical or practical enterprises. So, ordinarily, when we ask, for example, what the referent was of “acid” or of “water” before the development of modern chemistry, we are asking about the role of those terms, and of their reference-like relations, in explaining the reliability of chemical practice, *as that reliability improved throughout the history of chemistry*. So, we correctly assign, as the referents of those terms, kinds defined by “deep” or “important” factors, rather than by “superficial” or nominal ones. The appropriateness of such assignments depends, not on some extra-epistemic judgments of depth or importance, but on the nature of the explanatory questions which are typically raised in philosophical contexts. No non-epistemic judgments need be involved.

There is a sense, of course, in which epistemic judgments about referential hypotheses about natural kinds are related to a kind of normative judgment regarding the disciplinary matrices within which those terms are employed. The hypothesis,  $H$ , that  $t_1, \dots, t_i$ , refer within the context of a disciplinary matrix,  $M$ , to kinds with natures  $n_1, \dots, n_i$  implies a sort of hypothetical imperative “if you are to achieve what practitioners in  $M$  achieve, you must deploy conceptual resources which mirror the accommodation of conceptual resources to causal structures which  $H$  attributes to the uses within  $M$  of  $t_1, \dots, t_i$ .” Someone who advances  $H$  need not, however, in any way endorse the achievements of  $M$  as aims. A pacifist can offer a referential hypothesis about terms employed in military engineering without endorsing any of the practices of military engineers. She will merely be offering (part of) an explanation for their success in achieving aims she does not endorse.

Of course, what I claim on behalf of the naturalist moral realist is that just the same sort of treatment is appropriate when we consider referential hypotheses about moral terms and the natures of the entities to which they refer. In particular, I claim that there is no barrier in principle to the confirmation of such hypotheses by the hypothetical extraterrestrial anthropologist who, because of its psychological makeup, can have no commitment whatsoever to moral norms of practice. Let’s see, however, how well *our* usual reli-

ance on moral norms in evaluating hypotheses about the semantics of moral terms fits this picture.

2.5. *Achievements, Reliability and the Methodological Role of Informed Judgment.* Let's first consider why the judgments of chemists legitimately play an important methodological role in philosophers' investigations into the natures of chemical kinds. One part of the answer is obvious: chemists know a lot about chemical kinds. This is not analytic: there is no conceptual impossibility involved in supposing that the socially recognized experts on some subject know less, or at least little more, than do many others. Arguably, this was true of socially recognized medical practitioners until quite recently; it may still be true of experts in such domains as the psychology of gender differences. Nevertheless, in all but contrived cases, a term can't refer to a kind unless some people's use of the term reflects approximate knowledge of (at least some aspects of) the kind in question. When philosophers (or others) are able to identify people who are thus informed about kinds, they are (fallibly) justified in relying, with suitable caution, on those people's judgments about those kinds. That's our situation with respect to chemists and chemical kinds.

There is another aspect to our deference to chemists. Chemical kinds have natures which explain the ways in which the use of terms referring to them contribute to the achievements of *chemists* and of practitioners in related bodies of theoretical and applied practice. Now, I don't intend to provide a detailed analysis of the category *achievement*, but some things are clear about achievements. In the first place, not everything which is systematically and predictably brought about by some systematic human activities counts as an achievement of those activities, even if it is the sort of thing which would count as an achievement for some other body of practice. Achievements have to have something or other to do with the aims, purposes, interests, intentions, etc. of those practitioners whose achievements they are.

That something-or-other need not be a straightforward matter of bringing about what practitioners want or intend. Not only can we fail to achieve what we intend, we can achieve things we don't intend, even things we can't conceive of intending. [Consider early taxonomists in the Linnean tradition whose classificatory practices achieved a pretty good mirroring of the history of the evolution of life, even though most of them would have thought this a border-line incoherent project.] Similarly, groups of people working in concert can achieve things whose status as achievements (rather than mere predictable consequences) depends on the interests, aims, intentions, etc. of only a tiny minority of the group in question. Arguably this is the situation regarding the military and diplomatic accomplishments of almost all of the armed forces of modern nations.

Nevertheless, under many conditions, the judgments of practitioners regarding questions of methodology and practice provide a good *prima facie* indicator of which systematic consequences of their practice should count as achievements. When we defer to chemists regarding the natures of chemical kinds we are relying on their judgments as pretty accurate (though fallible) reflections of the "aims of chemistry" (as we might say).

Note that the ways we thus rely on judgments of chemists need not involve an endorsement of the aims of chemistry. Moreover, there is no reason in principle why reliable judgments about the natures of chemical kinds could not be obtained without any deference to chemists' judgments. The standard imaginary extraterrestrial anthropologists, if they included in their number their own experts on the properties of matter (*whether or not their articulation of that knowledge involved disciplinary categories or divisions of labor corresponding to our distinctions between, e.g., physics, chemistry and astronomy*), could in principle figure out what chemists were up to and what language-world relations contributed to their successes, without engaging in any such deference.

They would, of course, have to accord, to facts that they discerned about the judgments of chemists, the status of phenomena to be explained, but they would not have to defer to those judgments. Since reference involves epistemic access, they would be obliged to assign as referents for chemical terms entities with respect to which chemists practice afforded chemists *some* sort of epistemic success, but this constraint by no means entails that they should accept chemists' methodological judgments or chemists' assessments of their own epistemic successes.

Similarly, since the "aims" of chemistry—the parameters which discriminate between the accomplishments of chemical practice and its other predictable consequences—supervene in large measure on facts about the motivations and practices of chemists, the extraterrestrial anthropologists would have to take these factors into account in identifying the accomplishments of chemists. Here too, this requirement is much weaker indeed than the requirement that they defer to the judgments of chemists regarding the aims of chemistry.

Of course, if the current practice of philosophers in investigating the natures of chemical kinds is justified, then the conclusions reached by the extraterrestrial anthropologists would provide them with a justification for *adopting* the same pattern of deference to chemists which philosophers have adopted, but this does not indicate that their own methodology need have incorporated any norms distinctive to the human practice of chemistry. *Even when they were thus justified in adopting a pattern of deference to chemists in service of their anthropological work*, they would not be required by any principle of rationality to endorse the aims of chemistry. In so far as any norma-



tive judgments *followed from* the results of their inquiry (as opposed to methodologically underwriting that inquiry), they would have hypothetical form: “if you want to accomplish the sorts of things which human chemists accomplish, you need to use methods and employ schemes of classification similar to theirs in respect of such-and-such modes of accommodation to causal structures.”

The only norms which one would have to endorse in order to investigate in this way the nature of chemical kinds are the epistemic norms appropriate to evaluating hypotheses about the interaction of chemists and the world. Even these norms have—according to plausible naturalistic reliabilist conceptions of knowledge—a basically hypothetical form: if you want to (probably) find the (approximate) truth about ....’s you need to use methods....”<sup>8</sup>

Of course, I now urge that our reliance on our own moral judgments in assessing proposals about the natures of the (moral) good, justice, fairness, etc., or (equivalently) about the referential semantics of moral discourse, has just the same justification. We (let’s hope rightly) take ourselves to be informed moral practitioners, and we rely—with *serious qualifications*—on our own moral judgments as evidential with respect both to the extension of “good,” and the aims of moral practice. Theories of the natures of the good, etc. have, I suggest, just the same hypothetical normative import as do our theories of the natures of chemical kinds. “If you want to achieve the aims of moral practice, classify things this way:....” Thus I urge that a successful investigation into the metaphysics of morals could “in principle” be carried out by the hypothetical extraterrestrials and that neither they—nor we, if amorality is a psychological possibility for scientifically competent humans—would have to subscribe to, or accept, moral norms in order to carry out the investigation.

Of course these “in principle” counterfactual possibilities for completely morally unengaged investigations of the metaphysics of morals are not important to the actual, and actually appropriate, methods by which we do moral philosophy. They are merely an especially “philosophical” way of putting the point that the only normative commitments which one *must* make in order to explore the nature of the (moral) good are epistemic ones.

Still, these counterfactuals do make salient something important about actual moral philosophizing—both by professional philosophers and by others. Although it is neither possible nor necessary for most of us to conduct morally unengaged investigations of the metaphysics of morals, it is possible—and arguably it is desirable—for us, at some points in our lives, to accept our own moral commitments somewhat tentatively, while still assigning some evidential weight to our moral judgments, until we satisfy our-

---

<sup>8</sup> See Boyd 1985 and Miller 1987 for naturalistic treatments of epistemic norms for evaluating explanatory claims.

selves that the referents of “good” and similarly approbative moral terms are things we actually admire. It is this exercise to which we must turn if we are to have a adequate response to Marxist and other radical critiques of the social role of morality (about which more shortly). Such an investigative strategy, involving quite tentative commitment to principles previously unquestioningly accepted, will, I presume, be familiar to philosophers interested in the epistemology of religion as well as to those interested in the metaphysics and epistemology of morals.

What is important here, I believe, is that the level of skepticism about the *adoption* of moral norms which is compatible with an ordinary (non-extra-terrestrial, non-amoralist) investigation into the nature of the good far exceeds the level of skepticism about inductive or scientific methods which is compatible with the same investigation. It really is true that the only norms which one must *adopt* in such an investigation are epistemic ones.

Of course, I don’t claim that this is obviously the correct analysis of the relationship between moral judgements, on the one hand, and reasons for action or justifications for metaethical theories, on the other. If, for example, there is a deity of the sort Adams posits, then a full appreciation of the metaphysics of morals would involve acknowledging a non-hypothetical dimension of objective moral normativity. Thus a full defense of the position I am urging would require the defense of (something like) atheism. Still, since philosophical naturalists are committed to (something like) atheism anyway, I suggest that the role of normative moral judgments in theorizing about the nature of the good does not pose any special problem for those philosophical naturalists who are also moral realists. For them, as for moral philosophers generally, the methodology for investigating the metaphysics of morality is, as the early Rawls suggested, reflective equilibrium understood realistically (for a discussion see Boyd 1988), where the conceptions that are to be brought into equilibrium include, not only general and particular moral judgments and intuitions about the aims of morality, but also the best results of scientific, social scientific, *and metaphysical* inquiry. The fundamental disagreement between theistic and naturalist moral realists lies, not in the role of appeals to norms in their semantic theorizing, but in the metaphysical conceptions with which they require their moral and semantic theories to achieve equilibrium.

Thus, although Adams’ discussion of this matter raises absolutely crucial questions about normative judgments in semantic theory, and in the correlative metaphysical theory of the natures of things, there are answers to those questions which are fully adequate to the project of naturalist moral realism.

*2.6. Partial Denotation Revisited, I: Ideology and Partial Denotation.* I earlier indicated that I did not favor an appeal to the possibility of partial denotation

as a way of explaining the epistemic intuitions which underwrite Adams' notion of the critical stance. I do, however, want to insist that the notion of partial denotation can play a legitimate role in a naturalistic account of the semantics of moral language. In part, this is so because it is almost certainly a consequence of the role of social ideology in the determination of moral principles that moral terms partially denote.

I have not offered a detailed characterization of partial denotation and I will not attempt one here. Roughly, a term  $t$  partially denotes different kinds (relations, things, ...)  $k_1$  and  $k_2$  when  $t$  bears a strong reference-like connection to both  $k_1$  and  $k_2$  and when there is no other kind (or whatever)  $k^*$  such that the epistemic connection between uses of  $t$  and  $k^*$  better explains the contributions of the use of  $t$  to relevant achievements than do the epistemic connections between uses of  $t$  and  $k_1$  and  $k_2$ . To a good first approximation, the circumstances under which this happens fit into one of two scenarios (between which there is not always a sharp distinction):

1. *Scenario one:* There is a single disciplinary matrix  $M$  within which  $t$  is employed and the epistemic connection between the uses of  $t$  in  $M$  and  $k_1$  explains very nearly the same achievements in  $M$  as does the connection between  $t$  and  $k_2$ . In practice, practitioners in  $M$  do not distinguish between  $k_1$  and  $k_2$ ; their use of  $t$  corresponds, in a sense, to something like the union of  $k_1$  and  $k_2$ . Nevertheless, the reliability of their practice is compromised by this feature of their conceptual and linguistic practices. An improvement in reliability could be achieved by drawing the  $k_1$ - $k_2$  distinction and by replacing the existing use of  $t$  with the use of two terms (one of which might, but need not, be  $t$ ), one referring to  $k_1$  and the other to  $k_2$ . [This is *denotational refinement* in Field's sense (Field 1973).]

2. *Scenario two:* The epistemic connection between uses of  $t$  and  $k_1$  and uses of  $t$  and  $k_2$  are important to the explanation of quite different achievements, so that—depending on how you choose to individuate things—there are either two different disciplinary matrices involved, or two quite different families of achievements within the same broad disciplinary matrix. The distinction between  $k_1$  and  $k_2$  is not widely drawn so that the use of  $t$  corresponds, as before, to something like the union of  $k_1$  and  $k_2$ . Ordinarily (at least where the relevant achievements are overt rather than covert), cases which correspond to this second scenario will be ones in which the relevant achievements would be enhanced by denotational refinement.

The obvious example of the first scenario involves the use of the term "element" in chemistry before the distinction we now mark by the terms "element" and "isotope" was drawn. A plausible candidate for an example of the second scenario involves the term "species" as it is used in biology. A

number of biologists and philosophers have argued for “pluralism” about the species category: the theory that for different branches of biological inquiry—ecology, animal behavior, evolutionary biology, etc.—different notions of species are required, but that this need is unrecognized in practice, so that biologists work with a not-fully-adequate conception.<sup>9</sup> If this is true then the term “species” partially denotes each of several kinds of biological kinds in the way indicated in the second scenario.

What seems clear is that moral terms participate in something like this latter sort of partial denotation, at least on the optimistic assumption that they bear reference-like relations to properties (things, relations, etc.) which most of us would admire were the relations made explicit. The reason is that—as Marxists and others insist—there is a very deep relation of epistemic access between moral terms and various properties which are such that widespread admiration of them (in the case of the terms of approbation, contrariwise for terms of disapprobation) is a significant factor in explaining certain *achievements* of those classes (or other social groups, depending on which theory of exploitation and social inequality is the correct one) which occupy positions of disproportionate political power and wealth. There are, for example, systematic tendencies for the labor practices which are said to be “fair,” those wars which are said to be “just,” those character traits which are said to be “admirable,” those social aims which are said to be “compatible with the moral limitations of human nature,” etc. to be just the labor practices, wars, character traits, and social aims which serve the interest of those who rule, but not of those they rule. The epistemic access which the use of moral terms afford people to those properties plays an important role in explaining how it is that rulers *achieve* social stability and quiescence on the part of those they rule.

It is hard to see how one could deny that the relevant relations between moral terms and these ideologically relevant properties (relations, etc.) is reference-like. The distinction between partial denotation and less explanatorily central sorts of quasi-reference is not sharp, so I am happy to rest my case on the conclusion that moral terms participate in relations *very very* much like paradigm cases of partial denotation. It is important, however, to see that one especially attractive argument against the partial denotation interpretation fails.

It might seem plausible to hold that the *referential intentions* of participants in moral discourse decide the case in favor of reference (as opposed to partial denotation) to whatever ideologically less suspect denotata are available (the homeostatic cluster of components of human flourishing, or certain

---

<sup>9</sup> Actually several different but inadequately distinguished doctrines about species are all referred to as “pluralism.” I have picked the one which best illustrates the sort of partial denotation I have in mind. Wilson 1999 contains excellent discussions (and bibliography) regarding species pluralism in its various forms.

dimensions of resemblance to the deity, for example, in the case of “good”). We do not (most of us anyway) intend to refer to *convenience for the ruling class*, or whatever, when we use the term “good.”

There are two different reasons why this response is not decisive. In the first place, the whole point of theories of reference and natures which—like naturalists’ theories and like Adams’—reject the empiricist conception that reference is mediated by analytic definitions or otherwise conceptually privileged descriptive resources, is that we can and do refer to things such that we certainly don’t intend to refer to them under anything like the descriptions which in fact identify their true natures. If historical evidence indicated that 18th century biologists, in their use of species names, *intended not* (as opposed to *did not intend*) to refer to kinds with natural rather than supernatural natures that evidence would in no way undermine the claim that the natures of the species to which they referred are entirely natural.

The deeper point, in the case of the role of social ideology in moral and ethical discourse, is that the false consciousness on the part of participants in moral discourse, which is itself part of the relevant achievement of those in power, necessarily involves a sort of referential and intentional muddle. In the case of the word “good,” for example, there must be aspects of the ideologically determined, powerful-interest-serving, candidate referent with which users of the term identify. They must come to regard at least some aspects of that phenomenon with the same commitment and favorable attitude with which they regard features of the other, non-oppression-serving phenomenon to which, I presume, the term “good” also bears a reference-like relation. In so far as their actual referential intentions (*and* their allegiances) are concerned, both candidate referents are in the running.

*2.7. Partial Denotation Revisited, II: Reference and the Explanation of (the Possibility of) Future Achievements.* Earlier, in discussing why I did not think that an appeal to partial denotation would be the right response for the moral naturalist to the situations which Adams’ invokes to argue for (his understanding of) the critical stance, I suggested that a naturalist treatment of the semantics of moral discourse might require an appeal to partial denotation if it should turn out that—with respect to the morally relevant project of enhancing, through changes in social structure, the homeostatic unity of human goods—there are humanly possible future arrangements such that each is better by far (by any non-ideologically controlled moral standards) than our present arrangements, but which are such that no choice between them would be rationally dictated by norms that would be appropriate before the adoption of one or the other, and which are such that there is no way of “splitting the difference” between their moral virtues (see Boyd 1995).

Our discussion of reference and quasi-reference makes it possible for us to situate this (epistemic) possibility within the framework of the theory of reference. If one were interested *only* in the ways in which the use of chemical terminology contributed to the success of chemical practice before the discovery of isotopes, it would be *prima facie* appropriate to think of element names as referring to natures defined by the causal powers indicated by the positions of elements in the periodic table—that is by their atomic numbers—rather than to think of them as partially denoting. The standard conclusion of philosophers that element names partially denoted even before methods of classification sensitive to (what we now call the) element-isotope distinction were introduced reflects a particular (and perfectly legitimate) concern which philosophers usually have in seeing how the early practices of chemists put later chemists into a position to obtain the epistemic successes of contemporary (and future) chemistry. With respect to that future oriented-question, the relation of partial denotation between element names and, on the one hand, (what we now call) elements and, on the other hand, their various isotopes is explanatorily central.

The issue here is one of the individuation of disciplinary matrices. Philosophers of science are ordinarily concerned to formulate semantic hypotheses which offer explanations for the epistemic successes of disciplinary matrices in science *understood as ongoing enterprises* rather than hypotheses which (only) explain epistemic successes within historically delimited matrices. Just as philosophers interested in the foundations of science usually have such *future-oriented* semantic concerns, so do philosophers (and ordinary nonprofessionals) who are concerned with ethical and moral matters. Part of our usual aim in social and political morality is to look to the future of human social existence and to work towards the establishment of conditions even more conducive to human flourishing than those in which we live. The semantic theory (and thus the conception of the nature of the good and of other morally relevant properties) which underwrites an explanation of our potential for success in this future-oriented enterprise would have to recognize partial denotation for moral terms if there were two incompatible but morally incommensurable paths of moral progress open to us. It is by no means *obvious* (at least for philosophical naturalists) that this is not our actual position, so no naturalistic conception of ethics can reject this sort of possibility of partial denotation out of hand.

In fact, if by “moral realism” we understand the position that moral discourse must be held (semantically) accountable to achieving some sort of correspondence with reality, then acknowledging the epistemic possibility of this sort of partial denotation is constitutive of, rather than contrary to, the thrust of moral realism (Boyd 1988, 1995; for a general discussion of partial denotation as an aspect of reference see Boyd 1993, 2001b).

2.8. *Partial Denotation and the Grain of Truth in the Critical Stance.* Once we see that moral terms might partially denote different phenomena, each of which many of us would admire if we understood its nature, we can ask how one could rationally proceed if she were convinced that such partial denotation obtained. Here, I believe, we can see the grain of truth in Adams' discussion of the critical stance. Of course, faced with the recognition of partial denotation—especially of the latter future-oriented variety—one would not be rendered (rationally speaking) speechless. One could still indicate one's preference for one rather than the other of the competing moral schemes, and one could argue rationally for one's own preference. Of course, the sort of partial denotation which involved "moral ties," so to speak, between alternative moral futures would be such that one couldn't rationally persuade every fully rational moral agent if all the facts were known (otherwise there would be a determinate moral choice between the competing moral futures). Still, one might rationally persuade those with a similar preference structure to one's own (or others whose particular preference structures similarly dictated the choice one favored).

In this sense a critical stance would still be rationally available. Knowledge of the semantic facts regarding moral terms would no more settle the question "What is to be done?" for the morally committed agent, assuming that those terms partially denote in the way we are now considering, than knowledge about their determinate referents would settle that question for the extraterrestrial anthropologist or the human amoralist. There would still remain a domain for critical choice—a critical stance, if you like—an option to favor one rather than another moral future without making any mistake in rationally assessing the evidence.

The possibility of this sort of critical stance is important, but it does not tell against naturalist moral realism since that position, properly understood, anticipates and explains this sort of possibility. After all, it's not any particular surprise, even when it's disappointing, to learn that what seems to be a complex but unitary and coherent practical project actually involves us in unanticipated and difficult choices not settled by our original purposes. In fact Adams' theistic conception of the good avoids these problems only because of the *details* of his theology. If there were two equally loving gods, each with a slightly different (but loving) conception of our ends, or if the one god were imperfect (though loving) so that Her loving concern for us had ambiguous elements.... (you get the point).

I conclude that—absent an *independent* defense of something like Adams' theistic conception—we have no reason to acknowledge any sort of critical stance in the epistemology of morals whose explanation would pose difficulties for naturalist moral realism.

## Bibliography

- Adams, Robert. 1999. *Finite and Infinite Goods*. New York and Oxford: Oxford University Press.
- Block, Ned. 1980. "What is Functionalism," in Ned Block, ed. *Readings in Philosophy of Psychology*, vol. 1. Cambridge: Harvard University Press.
- Boyd, R. 1983. "On the Current Status of the Issue of Scientific Realism." *Erkenntnis* 19: 45-90.
- \_\_\_\_\_. 1985. "Observations, Explanatory Power, and Simplicity." In P. Achinstein and O. Hannaway, eds. *Observation, Experiment, and Hypothesis In Modern Physical Science*. Cambridge: MIT Press.
- \_\_\_\_\_. 1985a. "Lex Orendi est Lex Credendi." in Churchland and Hooker, eds. *Images of Science: Scientific Realism Versus Constructive Empiricism*. Chicago: University of Chicago Press.
- \_\_\_\_\_. 1988. "How to be a Moral Realist." in G. Sayre McCord, ed. *Moral Realism*. Ithaca: Cornell University Press.
- \_\_\_\_\_. 1989. "What Realism Implies and What It Does Not" *Dialectica*.
- \_\_\_\_\_. 1991. "Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds." *Philosophical Studies* 61: 127-148.
- \_\_\_\_\_. 1992. "Constructivism, Realism, and Philosophical Method." in John Earman, ed. *Inference, Explanation and Other Philosophical Frustrations*.
- \_\_\_\_\_. 1993. "Metaphor and Theory Change" (second version) in A. Ortony, ed. *Metaphor and Thought, 2nd Edition*. New York: Cambridge University Press.
- \_\_\_\_\_. 1995. "Postscript to 'How to be a Moral Realist'" in Paul K. Moser and J.D. Trout, eds. *Contemporary Materialism: A Reader*. New York: Routledge.
- \_\_\_\_\_. 1999a. "Homeostasis, Species, and Higher Taxa," in R. Wilson, ed. *Species: New Interdisciplinary Essays*. Cambridge: MIT Press.
- \_\_\_\_\_. 1999b. "Kinds, Complexity and Multiple Realization: Comments on Millikan's 'Historical Kinds and the Special Sciences'." *Philosophical Studies* 95 (1-2), 67-98.
- \_\_\_\_\_. 1999c. "Kinds as the "Workmanship of Men": Realism, Constructivism, and Natural Kinds." in Julian Nida-Rümelin, ed. *Rationalität, Realismus, Revision: Proceedings of the Third International Congress, Gesellschaft für Analytische Philosophie*. Berlin: de Gruyter.
- \_\_\_\_\_. 2001a. "Reference, (In)commensurability, and Meanings: Some (Perhaps) Unanticipated Complexities." in Paul Hoyningen-Huene and Howard Sankey, eds. *Incommensurability and Related Matters*. Dordrecht: Kluwer.



- \_\_\_\_\_. 2001b. "Truth through Thick and Thin," in Richard Schantz, ed. *What is Truth?*. Berlin and New York: Walter de Gruyter.
- Brink, David O. 1984. "Moral Realism and the Skeptical Argument from Disagreement and Queerness." *Australasian Journal of Philosophy* 62, 111-25.
- \_\_\_\_\_. 1989. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Field, H. 1973. "Theory Change and the Indeterminacy of Reference." *Journal of Philosophy* (70): 462-481.
- Goodman, N. 1973. *Fact Fiction and Forecast*, 3rd edition. Indianapolis and New York: Bobbs-Merrill.
- Kitcher, Philip. (1985). *Vaulting Ambition: Sociobiology and the Quest for Human Nature*. Cambridge, Mass.: MIT Press.
- Miller, Richard. 1985. "Ways of Moral Learning." *Philosophical Review* (94), 507-56.
- \_\_\_\_\_. 1987. *Fact and Method*. Princeton: Princeton University Press.
- Putnam, H. 1962. "The Analytic and the Synthetic." in H. Feigl and G. Maxwell, eds. *Minnesota Studies in the Philosophy of Science, III*. Minneapolis: University of Minnesota Press.
- Quine, W.V.O., and Ullian, J. S. 1978. *The Web of Belief*. New York: Random House.
- Shoemaker, S. 1980. "Causality and Properties." in P. van Inwagen, ed. *Time and Cause*. Dordrecht: D. Reidel.
- Sidelle, Alan. 1989. *Necessity, Essence, and Individuation: A Defense of Conventionalism*. Ithaca: Cornell University Press.
- Sturgeon, N. 1984. "Moral Explanations." in D. Copp and D. Zimmerman, eds., *Morality, Reason and Truth*. Totowa, N.J.: Rowman and Allanheld.
- Wilson, R. 1999. *Species: New Interdisciplinary Essays*. Cambridge: MIT Press.