Draft. Published in Hoyningen-Huhne and Sankey, eds. Incommensurability and Related Matters p 1-64 Springer 2001. ISBN 978-90-481-5709-9.

RICHARD N. BOYD

REFERENCE, (IN)COMMENSURABILITY, AND MEANINGS: SOME (PERHAPS) UNANTICIPATED COMPLEXITIES

Abstract: Received conceptions of the meanings of scientific terms assign to meanings an essentially benign methodological role: the meaning of a term consists of principles or inference rules which are, always or for the most part, (approximately) true or reliable. In fact, many scientific terms have meanings which are malignant: which are mainly false or misleading and which detract from, rather than contribute to, scientific progress. Kuhn's conception of incommensurability can be fruitfully extended to take account of malignant meanings. Malignant meanings are especially implicated in cases, like that of human sociobiology, in which the influence of social ideology on scientific practice is especially profound.

0. OVERVIEW

0.0. Malignant Meanings and the limits of Commensurability

In the present essay I propose a substantial revision of received naturalistic conception of the semantics of scientific (and other) terms and to the received understanding of the methodological role which the meanings of such terms play, especially with respect to questions of commensurability and incommensurability between competing paradigms or traditions of inquiry. Received naturalistic conceptions, in so far as they address questions of meaning at all, treat such meanings as *benign*: they assume that the beliefs and inference rules which constitute the meaning of a scientific term within a linguistic community will, typically, be approximately true or approximately reliable.

Such a conception, I argue, fails to recognize the ways in which the meanings of scientific terms can be profoundly misleading - indeed fundamentally incoherent - even when those terms are referentially unambiguous. The failure to appreciate the *malignant* aspects of meaning contributes, I shall argue, to a failure to appreciate important dimensions of incommensurability between research traditions. I propose an alternative conception of the meanings of scientific terms according to which such meanings are often considerably less *benign* than conceptions of meaning (whether naturalistic or not) usually assume, and I explore its relation to issues of real-life incommensurability between significantly different research traditions which share a common subject matter.

One conclusion I reach bears on the cogency of the view (contested by "postmodern" thinkers) that the practice of scientific methods can ordinarily be expected to contribute to intellectual progress through successively closer approximations to the truth. Here I propose to "split the difference" between modernist optimism and postmodernist pessimism. I argue that the grains of truth in traditional naturalistic semantic theories provide a justification for a significant level of optimism regarding progress by approximation, but that an appreciation of the role of *malignant* meaning dictates that we recognize the contrary tendency induced by incommensurability in what we may call *conceptual meaning*. In particular I suggest that when scientific practice is significantly influenced by social ideology - as it routinely is in the biological and social sciences - the (malignant) embedding of ideology in the conceptual meanings of scientific terms is often so substantial that the normal internal workings of scientific methodology prove insufficient to overcome the malignancy or to establish commensurability between mainstream scientific research traditions and those which are informed by ideological critiques.

0.1. The Kuhnian Background and the Standard Rebuttal

Kuhn (1970) introduced the issue of (what we may call) *methodological incommensurability* between competing scientific paradigms, an issue that others have raised about competing non-scientific enterprises, like moral or political conceptions. Kuhn's arguments for methodological incommensurability depend on the by now widely recognized theory-dependence of scientific methods. They depend as well on two much more controversial claims about semantics of scientific language, claims which posit two sorts of what we may call *semantic incommensurability* between competing paradigms. According to Kuhn, terms occurring in competing paradigms are *conceptually incommensurable*: they differ in what we might call *conceptual meaning* so that scientists from competing paradigms necessarily fail to communicate: they talk past one another. A consequence of this conceptual incommensurability, Kuhn says, is that there is *referential incommensurability* between the discourses of the two paradigms: the same terms when deployed in the two paradigms, because they have different meanings, must have different referents.

There is a near consensus in the philosophy of science that Kuhn profoundly exaggerates the extent of methodological incommensurability in the history of science and that his mistakes lie in his approach to the semantics of scientific terminology Other interpretations of Kuhn are possible which attribute to him less dramatic conceptions of semantic and methodological incommensurability (see, e.g., Hoyningen-Huene 1993). It is not my aim here to dispute them. Instead, I am concerned to argue that, even if Kuhn's technical arguments for incommensurability are as unpersuasive as I believe they are, the phenomena of semantic and methodological incommensurability to which he drew attention are even more widespread and important than he knew. I do not, however, advance my theses about malignant meanings as an exegesis of Kuhn's own views. To my knowledge he did not explore examples of the sort I discuss.. What we might call the *standard rebuttal* to Kuhn's arguments for incommensurability deploys a "causal" or "naturalistic" conception of reference to rebut the inference from conceptual incommensurability to referential incommensurability, by showing that terms associated with quite different conceptual resources can share a common referent and that this phenomenon is common in the actual history of science.

Naturalistic conceptions of reference are not, by themselves, sufficient to rebut claims of incommensurability for the sorts of cases Kuhn considers. Even when two conceptually divergent paradigms share a common subject matter (that is: are referentially commensurable) it still might be the case that the differences in conceptual meaning between the terms in the two paradigms are so great as to undermine the prospects for methodological commensurability. Although the standard rebuttal de-emphasizes issues of meaning in favor of issues of reference, it remains true, I'll argue, that all versions of the standard rebuttal are committed to a *benign* conception of the meanings of scientific terms. They reflect the estimate that, ordinarily, the beliefs and inferential and explanatory practices central enough to a paradigm or research tradition to count as features of the conceptual meanings of its terms will be approximately true enough (in the case of beliefs) or sufficiently reliable (in the case of methodological practices) enough that (a) they will contribute to the establishment of methodological commensurability with other paradigms with the same subject matter, and (b) most of them will be recognizable as insightful approximations in any successor paradigm.

0.2. Overview; Critique and Lessons

It is this benign conception of the meanings of scientific (and other) terms which I propose to criticize. Using examples from the emerging but highly influential tradition of research in "evolutionary psychology," I identify a class of inferential patterns or "scripts" connecting evolutionary and genetic premises to psychological conclusions which have the following properties:

1. They are profoundly unreliable.

2. Their unreliability is a logical consequence of fundamental and explicitly acknowledged theoretical principles in evolutionary psychology.

3. This unreliability is unrecognized.

4. Instead, these inferential scripts are central to the methodological practices of evolutionary psychology to such an extent that

    a. they form the basis for the central explanatory strategies in the tradition,

    b. acceptance of these practices - or an appreciation of a deep theoretical criticism of them - is a prerequisite to understanding the literature in the discipline, and

    c. criticisms of these inferential patterns are, for all practical purposes, unintelligible to practitioners in the tradition.

5. The contrast between these inferential practices and analogous practices in other disciplines which study human social behavior are such as to profoundly limit the possibilities for reciprocal criticism, and thus the prospects for methodological commensurability between the relevant disciplines.

6. Inference from similar past cases suggests that the resulting methodological incommensurability is unlikely to be resolved by developments internal to the relevant traditions. In cases like that of evolutionary psychology where social ideology is instrumental in

establishing patterns of inference it is often the case that external political criticism is a prerequisite to the establishment of a context in which commensurability is possible.

I argue that the inferential scripts in question are, in a perfectly good sense of the term - one central to the theory of communication - part of the *meanings* of the relevant terms in evolutionary psychology. They are, however, malignant rather than benign in so far as commensurability is concerned. I draw from this and other examples five lessons.

Lesson One Traditional empiricist conceptions of the meanings of general terms implied that their meanings consisted of exactly true beliefs about their referents. We need to go much further in rejecting this conception than we have thus far. In particular, the conceptions which are part of the *conceptual meaning* of a scientific (or other) term need not even reflect a coherent conception of its referent. Philosophers who (as they should) incorporate descriptivist elements into their naturalistic conceptions of reference need to take account of this phenomenon.

Lesson Two Kuhn was right to hold that features of the meanings of scientific terms can be an impediment to commensurability. The considerations which establish this point do not require any rejection of naturalistic conceptions of reference or any claims about referential incommensurability.

Lesson Three The sort of incommensurability which arises from differences in conceptual meanings is especially likely to arise in cases in which several relatively independent traditions examine the same subject matter, as they do in the case of inquiry into human social behavior and human psychological potential. The prospects for incommensurability are enhanced in such cases by the methodological impact of social ideology.

Lesson Four For this reason, the sort of meaning driven incommensurability we are considering is especially likely to arise in traditions of moral and political inquiry.

Lesson Five In many cases, the epistemological conceptions necessary for an understanding of how, and under what circumstances, commensurability can be established will constitute a *political epistemology*: one which assigns an epistemic role to some political developments external to the traditions in question. Even in cases in which issues of social ideology or of politics in the usual sense do not arise, it will characteristically be the case that the establishment of commensurability between initially incommensurable disciplines will be a matter of the internal politics of the relevant scientific communities and of the larger institutional structures within which they are embedded.

1. INTRODUCTION.

1.0. Kuhn on Semantic and Methodological Incommensurability

Kuhn's classic (1970) discussion of (alleged) incommensurability between consecutive paradigms is the *locus classicus* for discussions of the possibility that for some disagreements in science (and, by extension, in other domains) there might not exist methods which are both *rational* and *fair* (to the competing positions) for their resolutions. The possibility of this sort of thing, for the sciences at least, is established by the phenomenon (emphasized by Kuhn and Hanson (1958), but also by many logical empiricists in the 1950's and by scientific realists in the 1960's) of the *theory-dependence* of scientific methods.

Once it is recognized that the dictates of scientific methods are always determined by background theories it becomes clear that it is possible that there should be competing positions with theoretical commitments so different that the methods justifiable by the standards of both their positions would be insufficiently powerful to dictate a resolution of the dispute between them.

Kuhn, of course, asserts that such incommensurability is actual: that, for example, Newtonian mechanics, as it was formulated at the end of the 19th century, was incommensurable with the relativistic conception that replaced it.

Such incommensurability claims in the history of science have suggested to many thinkers that the same sort of incommensurability might obtain between consecutive states in other domains, like ethics or theology. Importantly, it is frequently suggested that the same sort of incommensurability also obtains with respect to scientific or other "paradigms" when these are not consecutive stages of a single research tradition, but rather different approaches to apparently the same subject matter arising in different cultural or intellectual traditions. Thus Kuhn's work has become integrated with broader traditions of relativism in anthropology and other social sciences. I'll be especially concerned with the possibility of this sort of incommensurability in the present paper, but it will be useful to begin by discussing the issue of the incommensurability of consecutive scientific paradigms.

In Kuhn's discussion of the transition between Newtonian and relativistic mechanics, an important appeal to *semantic incommensurability* enters into his defense of the *methodological incommensurability* of the two paradigms. It seems *prima facie* plausible that experiments and observations acceptable by Newtonian standards confirmed relativistic rather than Newtonian conceptions of physical magnitudes. Consider, for example, precise measurements of the orbit of Mercury or of the energy required in a cyclotron to accelerate particles as their velocity approaches the velocity of light. According to this conception, scientists, using fair methods, discovered that the laws of Newtonian mechanics provide only an approximate description of the relevant magnitudes and that relativistic mechanics provides a better approximation.

Kuhn rejects this understanding of the evidence: he denies, that, for example, Newtonianly acceptable measurements of mass confirm its velocity dependence, and he denies that Newtonian mechanics can be seen as an approximation to relativistic mechanics in the sense in question. Oddly, given his historical concerns, Kuhn's arguments for this position do not turn on subtle considerations about measurement techniques or other methodological practices of early 20th century physicists, or on a subtle discourse analysis of their papers, letters or public debates. Instead, they rest on a highly abstract conception (probably borrowed from Carnap - see e.g., Carnap 1950) of the semantics of theoretical terms. According to this conception, the most fundamental laws involving such a term constitute its *meaning* (in the sense of its *analytic* definition) and *determine its reference* in the sense of constituting an *exact* definite description of the phenomenon to which it refers.

On these grounds Kuhn claims that the meaning *and the referent* of the term "mass" changed during the transition from the Newtonian to the relativistic paradigm, since in the latter, a fundamental Newtonian law about mass (the conservation law) is not retained. Methodological incommensurability is assured because the competing paradigms involve incompatible meanings for their key terms *and do not even share a common subject matter of which their respective theories may be seen as providing approximate knowledge.*

1.1. Commentary: Semantic Incommensurability and the Dialectics of Inter-"Paradigm" Interaction

It will be important for our purposes to note that there are two different dimensions of semantic incommensurability which, according to Kuhn's arguments, obtain between competing consecutive paradigms and contribute to methodological incommensurability. The key terms in the two paradigms are said to differ in *meaning*, so that there is supposed to be *conceptual incommensurability* between consecutive paradigms. Thus defenders of the competing paradigms will fail to understand the conceptual resources deployed by their opponents.

There is supposed to be as well *referential incommensurability* between such paradigms: the key terms they have in common are supposed to have different referents in the usages of the two paradigms, even when this semantic fact is not evident to the parties to the interparadigm disputes. Thus, according to Kuhn, defenders of two such paradigms are "talking past" each other in two different ways: they fail to understand each other's conceptual resources and they are talking about different things even though they think their uses of the relevant terms are univocal.

Now, Kuhn thinks that conceptual incommensurability of the sort which he believes obtains between competing scientific paradigms entails referential incommensurability, but even if we accept his analysis of scientific paradigms for the sake of argument we can see that the two sorts of incommensurability are, in general, logically distinct. Thus for example when two people are discussing the very same acquaintance regarding whom they have very different experiences and beliefs, the differences in their presuppositions regarding her can be so great as to sharply limit their ability to understand one another, so that there can be conceptual incommensurability with respect to the (name of) the person in question even though there is no referential incommensurability.

Similarly, each of two discussants can be referring to a different one of two very similar people with the same name without being able to discern the ambiguity of the name - indeed without experiencing any conceptual discomfort at all. In such cases, there will be referential incommensurability without conceptual incommensurability.

Two points are important for our purposes about the ways in which these two sorts of incommensurability are related to the dialectics of the interaction between defenders of competing paradigms or research traditions. Consider first the impact of the dialectic of argumentation and persuasion between paradigms or research traditions on the two sorts of semantic incommensurability. Referential incommensurability is permanent: it cannot be eliminated or attenuated as a result of argumentation and persuasion. Two non-coreferential natural kind (magnitude...) terms cannot come to be coreferential without there being a change in the propositions expressed by the laws or generalizations in which they figure.

In the case of conceptual incommensurability, on the other hand, it is by no means obvious that this is so. Conceptual changes within two initially conceptually incommensurable traditions of inquiry - especially changes which result from the dialectics of argumentation between their respective practitioners - might reasonably be expected to establish conceptual commensurability, without any change in subject matter. Only if one accepts quite particular views about the reference of scientific terms like Kuhn's according to which (at least some kinds of) conceptual incommensurability for scientific terms entails referential incommensurability would one be inclined to deny this possibility.

If we turn our attention to the dialectics of methodological incommensurability a similar pattern emerges. When paradigms are referentially incommensurable, their permanent referential incommensurability entails a permanent methodological incommensurability, absent a revision in the referential semantics of the relevant vocabularies. In the case of conceptual incommensurability, methodological incommensurability is entailed - there cannot be methods which accomplish fair adjudication where there is not mutual intelligibility - but, precisely because conceptual incommensurability need not be permanent, neither must the methodological incommensurability it induces be permanent. Again, unless we adopt something like Kuhn's conception of the relation between conceptual and referential incommensurability, there is no reason to exclude the possibility that the dialectic of argumentation and persuasion between practitioners of two competing but conceptually incommensurable paradigms should result in developments within each which establish conceptual commensurability and, thus, eliminate a barrier to methodological commensurability.

It is important to note, of course, that even when conceptual commensurability is thus achieved there remains the apparent logical possibility that the (now semantically commensurable) paradigms might still diverge sufficiently that methodological commensurability is precluded. Still, the picture of initial conceptual (and thus methodological) incommensurability eventually overcome by the insights produced through the dialectics of scientific disputation and persuasion is *prima facie* familiar from, for example, the interactions which take place in the early stages of interdisciplinary investigations.

By contrast, the situation envisioned in Kuhn's treatment of the transition between Newtonian mechanics and special relativity, in which scientists believe that they have understood one another - and even award Nobel prizes for work which establishes new "paradigms" - only to learn from philosophers and historians that they have not even been studying the same phenomena, has, almost everyone agrees, never happened. It is perhaps surprising, therefore, that most philosophical discussions of commensurability and incommensurability in science have focused on the doubtful phenomenon of referential incommensurability rather than on the apparently real phenomena of conceptual incommensurability and the resulting (perhaps nonpermanent) methodological incommensurability. It is this inattention to real life conceptual incommensurability which I shall aim to criticize, and to begin to remedy, here.

1.2. The Standard Reply a la Quine and (Middle Period) Putnam. I: Reference as Primary

Since the publication of Kuhn 1970 a standard response has emerged in the philosophy of science which relies on causal (or "naturalistic") conceptions of reference of the sort initially introduced by Kripke 1971, 1972; and Putnam 1972, 1975 (but see also Feigl 1956) and which is influenced by Quinean critiques of analyticity and of related conceptions of meaning. Roughly, the idea behind the standard response is that the real semantic issue raised by Kuhn's examples of alleged incommensurability is one of reference rather than of meaning, and that - once the right theory of reference is employed - Kuhn's arguments for semantic incommensurability, and thus those for methodological incommensurability, fail. There is considerable variability in the details of presentations of what I am calling the standard reply, but I think that it is fair to say that what has emerged in their articulation are two basic ideas.

First, the referential relation between a word and a natural kind, property, magnitude, etc. is constituted by epistemically relevant causal relations between the use of the term in practice and the properties of its referent: relations which serve to establish a tendency for what is predicated of those terms to be approximately true of their referents.

Second, such a causal or naturalistic conception of reference for scientific terms makes a separate theory of meaning for such terms (as such theories are ordinarily understood) either inappropriate altogether or, at any rate, considerably less important for methodological issues like commensurability than has traditionally been thought.

What the versions I'll be interested in of the standard reply have in common include the following more specific claims about the relationship between meanings, true descriptions, and reference:

1. The epistemically relevant causal connections which establish and sustain reference will, at least in the case of theoretical terms in science, always depend in part on the prevalence within the appropriate community (perhaps just the community of experts) of some beliefs which are approximately true of the referents of the terms they use, but

2. nothing in the causal or naturalistic conception of reference requires that any of these beliefs be *exactly* true.

3. Furthermore, there is no place in the theory of the semantics of scientific terms for any notion of analytic truths about them. It may be possible to reconstruct philosophical intuitions about the "meanings" of scientific terms by treating as parts of the "meaning" of a scientific term at a given time the most fundamental of the laws or principles about it which are accepted in the relevant scientific community at that time. Still, these *law-cluster meanings* play no special *linguistic* or *semantic* role. In particular,

4. when the acceptance by the relevant community of approximately true beliefs is central to the establishment of reference for scientific terms (as it is in all typical cases), the beliefs in the term's law-cluster meaning need not play any special role. Ordinarily, lots of them will be approximately true, but there is no reason why this must be so. Nor is there any reason why approximately true beliefs which are *not* part of the meaning cluster cannot play as important a role in establishing the epistemically relevant causal connections as those in the cluster. Indeed, because of the importance to epistemic access of techniques of, and low-level beliefs about, measurement and detection, they typically will. Thus,

5. when two different communities use scientific terms (or other terms for which a causal or naturalistic understanding of reference is important) with the same referents, this fact need not be reflected in sameness or significant similarity of the law-cluster "meanings" of the relevant

terms in the two communities.

6. There is another semantic or linguistic role which traditional conceptions assign to meaning beyond grounding analytic truths and determining reference *via* exactly true descriptions. The meaning of a term may be thought of as something which a speaker must know in order to be linguistically competent with respect to the term. Law-cluster "meanings" aren't meanings in this sense either. In the first place, causal/naturalistic conceptions of reference all entail the possibility of "borrowing" linguistic and referential competence by deferring to experts or other knowledgeable people, so familiarity with the relevant law-cluster cannot be generally required for linguistic competence.

If we turn our attention to the referential competence of the relevant experts, it *will* be the case that that competence depends on the acceptance by experts of approximately true beliefs about the phenomena to which reference is achieved *and* that in the typical case many of the beliefs in the relevant law-cluster will be among those which contribute to the establishment or maintenance of reference in this way. Nevertheless, lots of other approximately true beliefs will play a role in maintaining reference and approximately true members of the law-cluster ordinarily play no distinguished role: their contribution to reference is the same as that of other relevantly approximately true and widely held beliefs. To think otherwise is to think that the beliefs which a community judges most central at a given time will always play some special role in establishing the referential semantics of their language. This is just what causal/naturalistic conceptions of reference deny: what determines reference are epistemically relevant causal relations and the history of science shows that these do not reliably correlate with conceptual centrality as judged by practitioners.

It is important to see that it is no part of the causal or naturalistic conception of reference to deny that *sometimes* the correct explanation for the divergence of conceptions or law-clusters between two paradigms or traditions of inquiry with apparently the same subject matter is that appearance differs from reality and they really lack a common subject matter. The point is, instead, that a divergence of the most central descriptions of that subject matter (a divergence in meaning in the "law cluster" sense of that term) need not dictate such a conclusion. Law-cluster "meanings" are, in an important sense, not linguistic or semantic phenomena: they aren't really *meanings*.

1.3. The Standard Reply, II: How Commensurability is Supposed to Obtain

When we turn to the question of how it is that methodological commensurability obtains (when it does obtain) between consecutive paradigms (and, by extension, in other cases of large-scale disagreement about a common subject matter) the standard reply *does (at least in the hands of Putnam 1962)* sometimes, perhaps, invoke the law-cluster meanings of scientific terms. Two competing conceptions or theories are commensurable just in case there are sufficiently many beliefs common to the two conceptions, and sufficiently many of them are relevantly approximately true, that they dictate reliable methods for resolving the disagreement between the two conceptions. Those methods thus held in common - when they are grounded in relevantly approximately true beliefs - provide the mechanisms by which actual properties of the common subject matter are indicated: they give the world a voice in the dispute, so to speak. The voice thus accorded to the phenomena in question explains how commensurability is achieved.

Once the view that the most fundamental laws within a paradigm or theoretical tradition are analytic is abandoned (as naturalistic and causal conceptions of reference require) the possibility becomes clear that empirical evidence will require the abandonment of one or more of these laws. Putnam 1962, in a classic (but very early) stage in the articulation of what I am calling the standard view, does however assign to such laws a special methodological status. He introduces the notion of a statement being *analytic at a time* when *at that time* it possesses the immunity from disconfirmation which genuinely analytic would *always* have, and he argues that the elements of a law-cluster have this property, even in the presence of what might seem to be disconfirmatory data, until adequate replacements have been proposed. Only when scientists have such replacements in mind do they consider rejecting fundamental laws.

Putnam 1962 assigns a distinctly semantic role to law-cluster meanings: a term retains its meaning/referent through a period of profound theoretical change just in case sufficiently many elements of its law-cluster are preserved. Importantly this semantic interpretation of law-clusters was (I think rightly) abandoned in later articulations of the standard reply (see, e.g., Putnam 1972, 1975; Field 1973; Boyd 1993, 1999a). Reference came to be seen as a matter of the right sort of causal relation between the use of a term and its referent. The special status of elements of the law-cluster came to be recognized as a methodological rather than a semantic or linguistic matter.

Let me explain. In the first place, laws acquire the status which Putnam calls analyticity at a time by being central to the scientific work of that time, This requires that they be laws of quite general scope (at least this is true of the examples Putnam gives) but they must also be laws which are highly well confirmed - otherwise they wouldn't have become so central to practice. Thus, when such a law appears to be challenged by anomalous data ordinary methodological considerations dictate that the challenge be assessed in the light of the very considerable body of evidence which favors it.

Putnam insists, however, that the laws in question cannot be overturned until a suitable replacement is proposed. In this regard, they might be thought to differ from other well confirmed laws or theories. Here too the requirement appears to be methodological rather than linguistic or semantic. In the first place, a well confirmed and important law will fit appropriately into the prevailing conception of what the relevant phenomena in nature are like (that's part of what it is to be well confirmed and important). When experimental data or other observations appear to compromise a law with these features, the alternative to rejecting the law is always to attribute the compromising observations to the operation of not-as-yet-understood factors, rather than to the falsity of the law itself. As anyone who has tested the predictions of Newtonian mechanics in a Freshman physics laboratory can testify, this is often the better choice.

I suggest that the requirement that an alternative be available which is comparable in scope with the law in question, and which can be comparably well integrated into the rest of science is, to some extent, a reflection of rational methodological standards for choosing between the alternatives just mentioned. Only when there is a credible conception of what the relevant phenomena are like which is compatible with the rejection of a law do we ordinarily have a good reason to attribute embarrassing data to its falsity rather than to the operation of unknown factors. This, too, is a methodological constraint rather than a linguistic or semantic one.

Finally, there is the point made by Putnam himself that fundamental and well established laws may be so methodologically central that scientists have no choice but to continue to apply them pending the acceptance of alternative laws, even when they seem to be challenged by recalcitrant data. Of course this is right; in fact often such laws continue to be applied as approximations after than have been disconfirmed and more accurate alternatives have been accepted. What is important for our purposes is that this, too, is a methodological rather than a linguistic or semantic constraint on the use of the terms appearing in such laws.

All these points are rendered more cogent by the observation that the sort of methodological entrenchment which Putnam 1962 attributes to the member of law-clusters obtains in the case of lots of important and well established scientific principles which are in no sense *fundamental* to the disciplines in which they are accepted. [Examples: the chemical formulae for everyday reagents, the classification of the *Felidae* in *Carnivora*, the currently accepted geological history of the Himalayas, the approximate values of physical constants as reported in your *CRC Tables*.] It is hard to see how there could be even a vague boundary separating the elements of law-cluster meanings from other well established and important principles and harder still to see what it could have to do with the semantics of scientific terms.

Thus the conclusion appears vindicated that the special epistemic or methodological role of law-cluster "meanings" is not a matter of their being *meanings* at all, and this is what later articulations of the standard response entail.

1.4. The Standard Response, III: Meanings as Methodologically Benign

I have characterized a standard response to the semantic presuppositions of Kuhn-style incommensurability arguments: one which involves denying that scientific terms have *meanings* in any strictly linguistic or semantic sense of the term and which portrays the possibility of commensurability between paradigms or research traditions as arising from shared *reference* of relevant terms together with an appropriate sort of overlap in the approximations to the truth reflected in the two paradigms or traditions. Since I propose to challenge this response, I want to be sure to represent it fairly. With this aim in mind, I want to articulate a more moderate position on meaning which I think would satisfy many defenders of the standard response. It, too, underwrites the conclusion that issues of meaning are not central to issues about commensurability and incommensurability and it too is, on the view I'll develop here, mistaken.

In the first place, fairness dictates that I acknowledge that philosophers responding to the incommensurability arguments of Kuhn had good reasons for focusing on the question of referential incommensurability. The arguments which Kuhn presented for referential incommensurability posed an extremely important challenge to prevailing conceptions of the referential semantics of theoretical terms. These conceptions were essentially *empiricist in origin and, hence, essentially descriptivist*: they assigned to the most fundamental laws involving a theoretical term precisely the role of uniquely picking out its referent. Thus when Kuhn deployed the same conception to defend the highly implausible conclusion that consecutive paradigms were always referentially incommensurable he produced a crisis in semantic theory requiring, as it turned out, the articulation of distinctly non-empiricist naturalistic conceptions of reference.

By contrast, there was no analogous crisis in the theory of conceptual meaning or, to put it more neutrally, the theory of mutual intelligibility across scientific "revolutions." In the first place the theory of "borrowed reference" which came free, so to speak, with naturalistic theories of reference made it clear that no special knowledge of the meaning of a scientific term is necessary in order for one to achieve mere referential competence with respect to it; it can be "borrowed" by linguistic deference to experts.

With respect to the question of conceptual intelligibility between experts, there was a similar absence of crisis. With respect to cases of "scientific revolutions" discussed by Kuhn - cases in which new "paradigms" are developed by practitioners trained in earlier ones and articulated using (what the founders take to be) conceptual resources shared with the earlier paradigms - almost any even remotely plausible conception of mutual intelligibility for experts, except one like Kuhn's which treats fundamental laws as analytic truths, will underwrite the (correct) conclusion that mutual intelligibility was largely achieved during the revolutions in question. Thus no new theory of conceptual meaning was required in order to articulate a response to positions like those of Kuhn.

In a word, it is just too easy to produce a theory which yields the conclusion that Copernicus, Galileo, Newton, Darwin and Einstein succeeded in communicating with other experts. No revolution in the theory of conceptual meaning is required. What I shall be arguing here is that the situation is quite different with respect to cases in which what is at issue is mutual conceptual intelligibility between relatively independent paradigms or traditions of inquiry with the same subject matter.

What I propose to do now is to articulate a more moderate position regarding the issue of conceptual meaning to which, I believe, many defenders of (versions of) the standard response tacitly subscribe and which, I believe, best captures the rationale for emphasizing issues of reference and de-emphasizing issues of conceptual meaning in the case of scientific terms.

The more moderate position I have in mind agrees with what I have called the standard response on two points. First, the establishment of the referential connection between scientific terms and their referents depends on all sorts of beliefs, methods and practices characteristic of (most) members of the scientific communities which use those terms. Second, although it is possible, in some or all cases, to rank such beliefs, methods and practices as more or less central or as more or less well established - as, for example, law-cluster conceptions of the meanings of scientific terms do - there appears to be no special way in which the most central or most highly confirmed beliefs play a special role in fixing the *referential* semantics of scientific terms of the sort which would distinguish their role as the role of *meanings*. What makes it a more moderate response is that it adds a hedge against the possibility that scientific terms really do have *conceptual* meanings: meanings which must be acknowledged in order to account for communication between scientists. Here are the components of the moderate position I have in mind:

1. Questions of conceptual meaning are questions about the prerequisites for mutually intelligible communication. If scientific terms have conceptual meanings, then the conceptual meaning of a term, t, within a paradigm or tradition of inquiry, P, involves a theoretical or methodological commitment, c, just in case a shared commitment to c is central to the possibility of intelligible communication between participants in P.

2. It is true that mutually intelligible professional communication within a paradigm (or between paradigms for that matter) depends on their being a sufficient level of similarity between the theoretical and methodological commitments of the relevant practitioners.

Nevertheless, it is unlikely that scientific terms have conceptual meanings in the sense just indicated. Probably all that is required for two scientists working within a paradigm to communicate is that there be some substantial similarity in their theoretical and methodological commitments, without there being any particular commitments that have the special status required of components of conceptual meanings.

3. Nevertheless, if it should turn out otherwise - if scientific terms do have conceptual meanings - then those meanings can ordinarily be expected to be *methodologically benign* at least in the mature sciences. That is: at least in mature sciences the fact that scientific terms have conceptual meanings can be expected ordinarily to contribute to, rather than to detract from, the prospects for methodological commensurability between paradigms or traditions of inquiry. Here's why.

a. In the first place, if scientific terms within paradigms or traditions of inquiry do have conceptual meanings, then successful communication between practitioners from different paradigms does not require that the conceptual meanings of the relevant terms be the same in the paradigms in question or even that they be mutually consistent. Conceptual meanings - whatever they may be - are not analytic definitions of scientific terms; scientific terms don't have analytic definitions. All that is required for communication between participants in different paradigms is that their adoptions of the conceptual meanings in their respective paradigms, together with whatever other conceptual resources they may possess, allow them to find each others theories and methods intelligible (even if profoundly mistaken). The dialectic of discussion and persuasion between paradigms can thus establish conceptual commensurability even when conceptual meanings differ.

b. Moreover, whatever conceptual meanings are they will surely reflect the best confirmed theories and best established methods in the relevant paradigms.

c. In mature sciences, at any rate, there is a general - although by no means absolute - tendency for the best confirmed theories to approximate the truth more closely over time and for methods to become more reliable (see, e.g., Boyd 1989, 1999a for a discussion of the dialectical relation between these two tendencies).

d. Thus it is to be expected that whatever theoretical principles or inferential practices are part of the conceptual meaning of a scientific term in a mature paradigm will tend to be approximately true or approximately reliable.

e. So, it is generally to be expected that when participants in two mature paradigms with the same subject matter subscribe to the conceptual meanings of the terms in their respective paradigms they will each be subscribing to approximately correct conceptions *of the same phenomena*. This will enhance, rather than diminish, the likelihood that their conceptual resources will be similar enough to underwrite conceptual commensurability. Thus conceptual meanings, if there are any, will ordinarily contribute to conceptual commensurability.

f. Similar considerations suggest that conceptual meanings should generally contribute to methodological commensurability in ways that go beyond their contribution to conceptual commensurability. Methodological commensurability obtains just in case there are sufficiently many points of (approximate) agreement between paradigms, among which are sufficiently many approximately true principles or approximately reliable methods, that these principles and methods underwrite fair and reliable methods for adjudicating the differences between them. Since the components of conceptual meanings will generally be instances of such approximate insights, they can generally be expected to make a positive contribution to the prospects for methodological commensurability.

g. Thus, if it should unexpectedly turn out that an account of the semantics of scientific terms *does* require identifying certain beliefs, practices or methods associated with such terms as components of their conceptual meanings, it will ordinarily turn out that the fact that scientists accept such meanings will be *methodologically benign* with respect to issues of commensurability between alternative paradigms or traditions of inquiry. Ordinarily, the beliefs and methods which will be part of the meaning of scientific terms within a paradigm will reflect real insights into its subject matter of a sort which will usually contribute to the establishment of conceptual and methodological commensurability with other approaches to the same subject matter.

So, differences in the meanings of terms will not ordinarily undermine the prospects for commensurability when two paradigms or traditions of inquiry share a common subject matter. In so far as semantic issues are at stake in matters of commensurability and incommensurability they are issues about reference rather than about meaning. Where referential commensurability obtains between paradigms in mature science we may *prima facie* expect both conceptual and methodological commensurability.

Since I propose to criticize this position I should indicate that I don't understand it to entail that there couldn't be referentially commensurable but conceptually incommensurable paradigms in mature science or that there could not be referentially and conceptually commensurable mature paradigms which resist the development of methods underwriting methodological commensurability. What I intend is for this defense of the *benign meaning thesis* to represent a highly plausible justification for the emphasis placed on referential commensurability when philosophers of science discuss the theses defended by Kuhn.

One other point is relevant here. It is widely agreed, I believe, that the sorts of epistemically relevant casual relations between the use of scientific terms and their referents which constitute the reference relation can obtain only when the relevant community has some significant approximately true beliefs about the referents in question. This is the grain of truth in the descriptivist conception of reference for scientific terms (see Boyd 1993 for a discussion). I suspect that some philosophers are attracted by the idea that the conceptual meanings of scientific terms will be given by approximately true fundamental principles within the relevant paradigms because they are attracted by a position even closer to descriptivism according to which the principles which are part of the meaning of a scientific term will always be among the approximate truths which help to establish the reference constituting causal relations. I believe that this is false - and indeed the arguments of the present essay are designed to show that it is profoundly false - but note that I have not built this particular sympathy towards descriptivism into the characterization of the benign meaning thesis.

1.5. Meaning and Communication in Practice

Traditional empiricist semantic theory assigns to the analytic definition, or "nominal essence" of a general term two distinct roles: the fixing of reference and the establishment of the possibility of unambiguous communication between speakers. The acceptance of the analytic definition of a general term is what determines linguistic competence with that term, and the descriptions which constitute its analytic definition define its referent or extension.

The new naturalistic conception of semantics which underwrites the standard response undermines both components of this conception, and it extends the critique so that it applies not just to analytic conceptions of meaning but to any accounts which assign to certain widely held beliefs (or other cognitive attitudes) a special role in determining referential competence with, or fixing the reference of, a scientific term. On the one hand, the possibility of "borrowed reference" and of deference to the semantic competence of experts entail that a speaker can use a term coreferentially with the rest of her community even if her beliefs about its referent are highly atypical. No deference to *conceptual meanings* - analytic or not - is required. Likewise, although some of the beliefs about a term's referent widely shared within a community may play a crucial role in determining its reference, these need not be components of its meaning - analytic or otherwise.

I agree with the naturalistic conception just sketched but I deny that it should be seen as supporting a benign conception of the meanings of scientific terms. I agree that research communities (or other communities of inquiry) can share a common subject matter even when their beliefs and other cognitive attitudes are very different, and that the possession by two such communities of a common subject matter ordinarily contributes to the prospects for methodological commensurability. I believe, however, that it happens fairly often that there are features of what we might call the *conceptual meanings* of scientific (or other) terms which differ between research traditions in such a way as to give rise to a quite robust form of methodological incommensurability even when the traditions in question unproblematically share a common subject matter. The sort of incommensurability which results from differences in conceptual meanings is, I shall argue, importantly different from that which might arise just from profound differences in theoretical conceptions. If I am right, what is involved in cases of incommensurability grounded in differences in conceptual meaning is really a matter of limits of mutual intelligibility between traditions - of "talking past" one another, just as Kuhn suggests - even when there is an indisputable commonality of subject matter. I'll argue, moreover, that in many such cases the conceptual meanings in (at least one of) the traditions will prove not to be benign. They will involve beliefs and inferential or explanatory practices which are so far from being true or appropriate that they detract from, rather than contribute to, the prospects for establishing commensurability.

Of course the phenomenon of talking past one another comes in degrees and grades. We are all familiar with situations in which two friends discussing a mutual acquaintance talk past one another until they realize that one of them knows and presupposes, whereas the other does not even know, that their acquaintance has recently married. Obviously cases like this do not raise questions about incommensurability.

Let me explain what I have in mind when I talk about conceptual meaning. Philosophers like Kuhn defend a law-cluster conception of the meanings of theoretical terms. As we have seen, they can be thought of as defending two related semantic claims about the role of fundamental laws in determining the semantics of such terms. On the one hand, they can be thought of as making a claim about how the reference of theoretical terms is fixed - as maintaining (mistakenly as it happens) that the referent of a theoretical term must be such that all or most of the central laws about it are (perhaps approximately) true. Alternatively, they can be thought of as making a claim about communication and mutual intelligibility. According to the second sort of claim, intelligible communication between researchers is impossible unless they accept the same law-clusters for the theoretical terms they employ.

I think that this latter claim is false, but I think that something like it is true, and the notion of conceptual meaning is supposed to capture the relevant grain of truth. In order to get a better understanding of what this grain of truth consists in, it is important to see some ways in which the law-cluster conception - understood as a theory of the prerequisites for communication - is *not* refuted. In the first place, it is not refuted by the phenomenon of borrowed reference. Suppose that a law-cluster theorist maintains that the law of conservation of mass-energy is part of the conceptual meaning of the notions of mass and of energy. She will be maintaining that acceptance of this laws is necessary for communication among physicists and others who deploy the terms "mass" and "energy" in doing science. She need not deny, for example, that a physics teacher can understand the error made by a student who misunderstands a thermodynamics lecture and writes on an examination that in closed systems mass is conserved but total energy decreases over time.

More importantly, the law-cluster conception of conceptual meaning is not undermined by the fact that some researchers - historians of science for example - can, by suitably immersing themselves in traditions with different law-clusters, come to understand, and compare, the doxastic and methodological commitments of the two traditions. Provided that this sort of specialist's understanding is not what underwrites the communicative capacities of the overwhelming number of participants in the traditions in question, the law-cluster conception of conceptual meaning would not be compromised.

Let's introduce some terminology here. Let's say that someone *engages with* some features or other of the conceptual or inferential resources of a tradition of inquiry just in case either (a) she completely accepts those conceptual resources and participates in the relevant inferences (let's call this sort of case *uncritical engagement*) or (b) she fully appreciates the relevant conceptual and inferential resources but engages in a serious theoretical or methodological critique of them, either rejecting some or all of them or adopting some intermediate skeptical or agnostic attitude towards them (call this *critical engagement*).

Using this terminology we may say that the law-cluster conception defended by Kuhn holds that the conceptual meanings of scientific terms are such that communication between scientists across research traditions ordinarily depends on their being *uncritically* engaged with the law-cluster associated with the relevant terms in the two traditions, *so that* - except for historically or philosophically sophisticated specialists who are critically engaged with their respective traditions - communication is impossible unless scientists accept the same law-clusters.

Of course this is mistaken, as the case of the transition between Newtonian mechanics and special relativity indicates, and we may use the terminology just introduced to identify the crucial error. Under the circumstances involved in this transition - in which the predecessor paradigm and the successor paradigm are both approximations to the truth and (more importantly) the claims made by the theories central to the two paradigms are *in ways obvious to practitioners in either paradigm* approximations to each other - critical engagement does not require specialized skills or knowledge (historical or philosophical skills for example) beyond those which are available to competent investigators in either paradigm. Thus, although critical engagement on the part of participants in the two paradigms *is* required to establish conceptual commensurability, this is no barrier to communication or to the establishment of methodological commensurability.

Thus Kuhn, for example, greatly exaggerates the conceptual barriers to communication between participants in the scientific "revolutions" he discusses. These "revolutions" were, after all, initiated by participants in the "paradigms" whose replacement they urged - participants who deployed the conceptual resources they shared with other participants in the earlier paradigm to argue for the new one. What I want to argue here is that, in other respects, Kuhn (at least in the key semantic argument on pages 101-102 of Kuhn 1970) greatly *underestimated* the range of conceptual resources which are properly counted as components of the conceptual meanings of scientific terms.

1.6. Engagement and Conceptual Meanings

I suggest that the notion of conceptual meaning can be usefully explicated in terms of the notion of engagement we have just deployed. Let's say that a law, doctrine, or inferential practice is part of the conceptual meaning of a term within a paradigm or tradition of inquiry just in case engagement with it (either critical or uncritical) is central to an investigator's capacity to intelligibly appreciate and discuss the findings and methods of the paradigm or tradition in question.

Of course, centrality is a matter of degree, so it would be open to someone to raise the (essentially Quinean) challenge that there is so smooth a "continuum" between the most and the least central doctrines and practices that the notion of conceptual meaning just defined fails to correspond to anything fundamental to the theory of scientific communication. This is an empirical issue, of course, so I'll provide examples of doctrines and practices which, I think the evidence will show, are presupposed in the literature in particular research traditions, and in the dialectic of argumentation and persuasion within them - and between those traditions and rival ones - *and* which are such that only someone who either assimilated those doctrines and methods as a member of the tradition in question or critically engaged with them from an alternative research perspective could make any scientific or methodological sense of the work and findings of the tradition in question. That there may be a continuum along conceptually or methodologically relevant dimensions connecting such cases to principles or practices not thus central to communication and understanding does not mitigate the conceptual centrality of the former.

This conception of conceptual meaning is, it should be noted, broader than that to which Kuhn appeals in arguing for methodological incommensurability. That conception - borrowed as it is from Carnap and other late logical positivists - makes the doctrines which are part of the meaning of scientific terms *analytic truths* with the consequence that obviously non-analytic doctrines (or methodological practices which are obviously not analytically justifiable) are not candidates. On the conception I am proposing, features of "paradigms" as Kuhn understands them - inferential practices, for example, or standards for experimental design, or non-"fundamental" but widely applied truisms - are potential candidates for conceptual meaninghood even when it is obvious that they are not analytic or analytically justified.

One special case of such features will illustrate the sort of conceptual meanings I have in mind and will also aid in our understanding of malignant conceptual meanings when we come to discuss them. Kuhn (1970) remarks that scientists working within a paradigm have "quasi-metaphysical" knowledge which allows them to know in advance, in typical cases, what the form of the solution to a scientific question will be. The phenomenon he has in mind is a matter of *projectibility* judgments in the sense of Goodman (1973). These are paradigm (or tradition, or whatever) dependent judgments of theoretical plausibility and they set the research agendas for paradigms. To a good first approximation, methodology within a paradigm involves (theory dependent) testing of *projectible* answers to scientific questions, subject to the important constraint that the test procedures control for those possibilities for experimental artifacts (or artifacts of observational technology) which are themselves suggested by *projectible* conceptions of the relevant experimental or observational situation (for a systematic treatment see Boyd 1985a, 1985b, 1990a).

Now, as Kuhn recognizes, coming to be able to make appropriate projectibility judgments is not entirely a matter of having explicit knowledge of the relevant paradigmatic theories. A large part of this sort of methodological sophistication involves tacit knowledge of inferential standards and practices which is acquired in the process of graduate education and through other methods of professional indoctrination.

What is important for our purposes is that engagement with at least the most important (tacit as well as explicit) standards of projectibility within a paradigm is a prerequisite to understanding its published literature or to engaging in scientifically or methodologically fruitful conversations with its practitioners about their work.

To see this, just try reading professional journals in some discipline with whose basic terminology you are familiar but in which you have not been (officially or unofficially) trained, and then share your impressions of the papers you have read with a sophisticated practitioner. You will find - and I am here reporting the experience of everyone in the philosophy, history or social studies of science - that you have missed, or misjudged, most of the important substantive or methodological issues.

The explanation, of course, will be that the authors of the papers in question will have designed their studies and written up the results within the framework of projectibility judgments dictated by the current state of the art in their paradigm or tradition. Their papers will provide evidence which bears on the choice between (what they and their readers will take to be) projectible answers to the relevant questions and they will have incorporated controls for artifacts in ways similarly dependent on projectibility judgments. They will presume that their readers are equally immersed in that paradigm and share basically the same methodological standards, including standards of projectibility.

Partly (but only partly) for that reason, they will not make explicit many of the central methodological judgements (of projectibility, in particular) which inform their studies. The other reason, of course, is that they could not make explicit most of these judgments if they wanted to. I do not mean merely that they could not classify these judgments as, for example, *projectibility* judgments, because they aren't familiar with Goodman's notion of projectibility. I mean instead that (as Kuhn notices - this is part of the importance of *exemplars* in determining the direction of research within a paradigm) most of those judgments are, as a matter of fact, only *implicitly* represented in the theories and practices of the paradigm or tradition. Moreover this is not a state of affairs which practitioners could fully remedy if they wanted to. As historians, sociologists and philosophers of science have discovered, the task of (even partly) explicating the methodological presuppositions of a paradigm or research tradition requires special skills which are not ordinarily part of the intellectual equipment of its practitioners.

Thus when you read the literature in the paradigm in question, and when you discuss scientific and methodological issues with its practitioners you will *always* fail to appreciate many of the fundamental substantive and methodological issues which are at stake unless and until you have (perhaps critically) engaged with those of its doctrines and practices which underwrite projectibility judgments. Thus the most central of these doctrines and practices must be counted as components of the conceptual meaning of the relevant scientific terms within the paradigm or tradition in question.

Similar considerations indicate that certain standard features of experimental design or of instrumentation - ones an understanding of which is presupposed in the relevant literature - will count as components of the conceptual meanings of those terms whose applications, in routine scientific practice, they help to determine. Without (perhaps critical) engagement with these features of paradigmatic practice you would be unable to understand papers in the literature or appreciate the dimensions of methodological or substantive argumentation within the paradigm.

Note that the account of conceptual meaning offered here accommodates exactly Kuhn's understanding of the importance, for a scientist's understanding of scientific concepts, of her immersion in her paradigm's professional practice and of her (partly tacit) appreciation of its exemplary research techniques and findings. The present account differs from Kuhn's more Carnapian conception of meaning for scientific terms, as it is deployed in his argument for referential incommensurability, in that it does not require of components of the conceptual meaning of a scientific term that they be either analytic (in the case of doctrines) or analytically justifiable (in the case of methods), nor does it require of each such component that it play a crucial role in determining the referent of the term. Only their role in underwriting intelligible communication among specialists distinguishes components of conceptual meaning in the sense proposed here from other features of the use of a term.

For this reason, the present conception can recognize as components of a term's meanings features of its deployment within a paradigm which are central to the paradigm's (communicative and experimental) practices even when they are not candidates for components of its meaning in Kuhn's more Carnapian sense. The conception offered here is, therefore, more consonant with what he had to say about conceptual understanding within a paradigm than is the Carnapian conception upon which his own arguments for referential incommensurability rest.

1.7. Conceptual Meanings and Conceptual Commensurability: Conceptual Commensurability as a Cognitive Achievement

Let's consider the question of conceptual commensurability between paradigms and research traditions in the light of the proposed account of conceptual meanings for scientific terms. One important historiographic point made by Kuhn (1970) is that coming to understand the doctrines and methods of a paradigm different from the one in which one has been trained is a significant cognitive achievement, often a very difficult one. It is this achievement which, among other things, permits us to avoid the sort of whiggish history of science which sees earlier scientific theories and practices as irrational in so far as they differ from our own.

The Carnapian conception of meaning upon which Kuhn explicitly relies permits us to see some of the ways in which this is true. In so far as fundamental laws are part of the meanings of scientific terms, then we can see why appreciating the rationality and achievements of paradigms other than one's own requires a special sort of engagement with unfamiliar conceptual material and the deployment of skills which are not ordinarily required of participants in any particular scientific paradigm (and which may be, in some sense, the special province of historians and philosophers of science).

Nevertheless, because Kuhn's Carnapian conception must do "double duty" as an account of analytic reference fixing descriptions, it cannot be deployed to explain another *quite Kuhnian* phenomenon. The achievement of conceptual commensurability between different research traditions is a significant achievement - requiring the deployment of skills and resources not ordinarily required for practice within a research tradition - even in cases in which the most fundamental laws accepted within the traditions are the same, and in which there is no serious prospect of a persuasive argument for either referential *or* methodological incommensurability. It requires, as we have seen engagement (critical or otherwise) with a complex body of partly tacit doctrines and inferential practices.

Of course the notion of conceptual meaning proposed here is designed precisely to explain this phenomenon. Methodologically or substantively significant conversation between research traditions always does require the *achievement* of conceptual commensurability through mutual engagement. When research traditions disagree, or even appear to disagree, in their doctrines or methods methodological commensurability cannot be achieved without the establishment of conceptual commensurability. There is never - or almost never - in such cases *immediate* methodological commensurability, and the reason that this is so *is* that scientists in different traditions use scientific terms with different meanings. Kuhn was right about that and, once this insight is distinguished - as it is here - from an empiricist conception of the relation between conceptual meanings and reference fixing, we can see that Kuhn's insight extends to a far greater range of cases than those which fit his model of "scientific revolutions."

It remains to examine the relation between conceptual meanings, properly understood, and methodological commensurability. Given that differences in the conceptual meanings of scientific terms within a paradigm or tradition will ordinarily be an *initial* barrier to the achievement of conceptual commensurability (and, thus to methodological commensurability) the question remains whether or not the acceptance of those meanings for terms by participants within a research tradition can ordinarily be expected to contribute to their ability to achieve conceptual commensurability with another tradition sharing a common subject matter. According to the *benign meaning thesis* the answer is "yes": ordinarily the components of meaning for a scientific term within a mature paradigm or research tradition will reflect substantial insights into its subject matter and these insights will facilitate a participant's understanding of the (approximately equally) genuine insights reflected in the conceptual meanings of the same terms in another mature tradition.

In the simplest case, on the benign view, participants in two such traditions would be able to rely on the insights of their respective traditions to come to a (perhaps critical) engagement with each other's traditions through exploration of each other's research literature and through serious substantive and methodological discussions. In more complex cases, the differences between the traditions might be such that even the establishment of conceptual commensurability might require experimental observational studies of the relevant phenomena conducted in the light of the special need to clarify points of tension between the traditions. But, according to the benign meaning thesis, even in cases of the latter sort the components of term meaning within either of the traditions can ordinarily be expected to reflect genuine (perhaps partial and approximate) knowledge of the common subject matter, and thus to contribute to the success of the relevant empirical studies.

This thesis is plausible enough if we focus our attention on the examples of "scientific revolutions" discussed by Kuhn. It is also rendered plausible by the reasonable idea that the doctrines and methods which are best established as parts of the conceptual machinery of either one of the two traditions are likely to occupy that position because they are among its best confirmed findings, and are thus likely to reflect real (approximate) knowledge of their common subject matter. It may also, as I suggest earlier, gain some plausibility from the conception that empiricists descriptivist conceptions of reference are right in one respect: that the conceptual meanings of scientific terms ordinarily contribute to the establishment of the reference relation for those terms by reflecting approximate knowledge of the phenomena to which those terms refer.

It is this benign conception of conceptual meanings that I propose to dispute. I think that there are scientifically (and socially) important cases in which, far from representing approximate knowledge or reliable methods, the conceptual meanings of terms within scientific research traditions are not only profoundly mistaken but, if made explicit, profoundly incoherent. For practitioners in such traditions, their acceptance of such meanings profoundly detracts from - rather than contributes to - the possibility of conceptual commensurability with other traditions sharing the same subject matter.

The existence of such cases, I shall argue, indicates that Kuhn was right to think that divergence in meaning can significantly compromise the prospects for methodological commensurability, even though he was mistaken in thinking that they ordinarily imply divergence in reference. The almost exclusive focus in the post-Kuhnian literature on issues of referential commensurability has thus led to an under appreciation of an important dimension of methodological commensurability.

Equally important are the implications of such cases for theories of reference. Conceptual meanings of scientific terms are, of course, important to the establishment of reference: they are central to the possibility of communication within relevant research traditions (this is how *conceptual meaning* is defined) and, since reference is a matter of *socially coordinated epistemic access* (Boyd 1993, 1999a), the establishment of social communication is central to the phenomenon of reference. It is also true that the epistemic access characteristic of reference is ordinarily achieved only when research communities have lots of approximately true beliefs and lots of approximately reliable methods. *But*, the contribution of conceptual meanings to reference need not be the one suggested by venerable descriptivist conceptions of reference: the conceptual meanings of scientific terms within a tradition need not provide an even approximately accurate conception of their reference nor even approximately reliable methods for studying them. Conceptual meanings can be - and in important cases are - utterly misleading and fundamentally incoherent. We now turn our attention to some examples of this phenomenon.

2. EVOLVED BEHAVIORS AND HUMAN NATURE: AN EXAMPLE OF INCOMMENSURABILITY IN THE MAKING

2.0. Cultural Studies (and Cognitive Science) vs. Sociobiology

A wide variety of disciplines in the humanities and social sciences - especially history, anthropology, social psychology, political science, economics and philosophy - have been concerned to provide explanations for large scale features of human social behavior: family structures, social arrangements, cooperation and competition, religion, warfare, divisions of labor, and the like. With the publication of Wilson 1975 there came to be recognized an alternative research tradition, initially called "sociobiology" and then later renamed (to avoid the embarrassment occasioned by obvious political excesses in its earlier stages) "evolutionary psychology." The aim of this new approach is to substantially reduce questions in human social psychology to questions in evolutionary biology and, thereby, to obtain significant insights into the patterns of human social behavior previously in the domain of the humanistic and social scientific disciplines just mentioned. The basic strategy is to deploy what are called "optimality models" in evolutionary biology to make estimates about probable patterns of behavior in early human societies and to infer, from these patterns, generalizations about human motivational structure and social psychology. It is within this research tradition that I propose to identify a family of terms with conceptual meanings which are malignant rather than benign. In particular I shall argue that there are inferential practices at work in contemporary evolutionary biology which:

a. Are profoundly unreliable,

b. reflect - if made explicit - an essentially incoherent conception of the evolution of human psychology, ones inconsistent in many points with explicit (and well established) doctrines established within the same research tradition,

c. are parts of the conceptual meanings of the key terms, and

d. render almost unintelligible - for those uncritically engaged with them - deep and correct methodological criticisms of evolutionary psychology arising from the other disciplines mentioned, so that

f. the prospects for methodological commensurability are greatly reduced.

[Wilson 1975 and the more popular Wilson 1978, together with Barash 1979 and Alexander 1979 provide an introduction to the early work in this literature, when the term "sociobiology" was standard. Betzig 1997 and Simpson and Kenrick 1997 provide a sample of the very large contemporary literature. Pinker 1996 provides an example of a recent popularization. Sherman and Reeve 1997 provides a methodological discussion of extraordinary sophistication. The key Journals include *Ethology and Sociobiology, Behavioral Ecology and Sociobiology, Journal of Social and Biological Structures, Human Nature, Trends in Ecology and Evolution*.]

2.1. Anecdotes

My confidence in the position I am going to defend here about conceptual meanings in evolutionary psychology has a history which provides some of the evidence I'll offer for it. For the last decade and a half I have been teaching an undergraduate course, "Science and Human Nature," which examines substantive and methodological controversies about human sociobiology. With the help of colleagues who have taught with me (Professors Nicholas Sturgeon and Karen Jones) I came to see that - as I'll indicate below - certain central inferential practices in human sociobiology are obviously in conflict with extremely secure findings in post-behaviorist psychology, findings which, in other contexts, human sociobiologists all acknowledge.

When I and my colleagues tried to explain this conflict to our students, and to criticize the inferential practices in question, what we found was that - especially in the cases of students who had previously studied evolutionary psychology, but even for those who had only taken introductory courses in evolutionary theory and animal behavior - it was extremely difficult for students to remember the criticisms from one lecture to the next or to apply them to actual cases of sociobiological inferences.

I do not mean merely that not all of our students were convinced by the criticisms we offered. That was to be expected. What I mean is that many of them reported the experience of finding the criticisms plausible sounding at first, but then finding themselves unable even to paraphrase them. When we would assign exercises in which the criticisms were made explicit (in short, logically uncomplicated, sentences) and students were asked to say how someone who accepted these criticisms would respond to very simple sample cases of the relevant sociobiological inferences, they often were unable to do so, and they themselves found this puzzling.

At first, we attributed this phenomenon to the methodological immaturity of undergraduate students, but our experiences, and those of other philosophers who presented the same criticisms to sophisticated professional human sociobiologists, persuaded us that something deeper was going on which rendered the criticisms extremely difficult for those initiated into standard patterns of evolutionary thinking about human psychology to understand, even though (as you shall see) the logical structure of the criticisms is hardly daunting.

The account of conceptual meanings presented here is my attempt to explain the phenomenon just described and to set it in an epistemologically and semantically informative context.

2.2. The Standard Pattern of Extrapolative Sociobiological Inference, I: Inferring from Evolutionary Scenarios to Motives

Terminology first: I distinguish between two sorts of projects in human sociobiology. The first project - let's call it the *explanatory* project - seeks evolutionary explanations for *independently identified* features of human psychology. It is a special case of a standard pattern of explanation in evolutionary theory in which evolutionary explanations are offered for independently identified phenotypic traits of organisms. There are controversies about the methods proper to such explanations (see, e.g., Gould and Lewontin 1979 for a famous criticism of some standard methods), but I will not be concerned with this project or those criticisms here. What I am concerned to discuss is the project of *extrapolative* human sociobiology (or extrapolative "evolutionary psychology"): the project of using the resources of evolutionary theory to determine which theories of human psychology are likely to be true, rather than to explain independently confirmed psychological theories.

There is an inferential pattern which is more or less characteristic of the extrapolative project just discussed. An evolutionary scenario regarding the evolution of certain behaviors, B, in early humans is presented. Usually, but not always, this scenario is said to have been retrodicted from evolutionary theory according to the dictates of an "optimality model." An underlying motivation, M, for B is then specified (with greater or less precision depending on the extrapolative argument in question) and evolutionary theory is said to "predict" or to "suggest" or otherwise to confer a privileged epistemic status on an hypothesis, H_M to the effect that M is an innate and relatively non-malleable feature of human psychology available for explaining human behavior under conditions different from those in which prevailed in the environment in which human psychology evolved Finally, it is my aim here to argue that standard inference patterns in evolutionary psychology are *malignant* components of the meanings of relevant terms, in the sense discussed above. I did not want to stack the cards in my favor by focusing on an inferential strategy which is even less credible that the one I discuss at length. It is widely recognized that there are deep problems with positing optimality everywhere (see Cosmides and Tooby 1987). In particular, if it turned out that animals routinely exhibited near optimal adaptations to environments quite different from those in which they evolved, then Darwinian evolutionary theory would have to be abandoned in favor of (something like) a theory of adaptation by design! [Sherman and Reeve 1997 offer an even more sophisticated taxonomy of inference patterns in evolutionary psychology. They consider applications of the "optimality everywhere" strategy in those special cases in which a pattern of behavior, characterized in what I have been calling *ecological parameters* (see Section 2.3.3), might have been subject to constant positive natural selection up until the very recent past. The considerations rehearsed in Section 2.3.3 raise doubts about whether constant selection for an *ecologically characterized* pattern of behavior must be reflected in the sort of constancy of selection at the *psychological* level which would be presupposed by the "optimality everywhere" inferential strategy in such cases, but this is an issue beyond the scope of the present paper.] (the *environment of evolutionary adaptation*; henceforth: *EEA*).

There is an important assumption which underwrites this inference pattern and which I will not challenge here. According to the "no one here but us hunter gatherers" thesis, although human populations have been subject to the operation of selective forces at all times from the development of the first complex agricultural societies to the present, since humans stopped living in small hunter gatherer groups there has been no uniformity to the direction of selection with respect to psychological characteristics. Because the environments in which humans have lived since the hunter gatherer experience have been so varied, and have changed so rapidly, psychological traits which were favorable for reproductive fitness under one social arrangement were likely to be unfavorable under subsequent ones and *vice versa*. The effect of this rapid variation in selection pressures is that the basic developmental psychology of humans has not changed significantly since we were hunter gatherers. The EEA for *contemporary* human developmental psychology was the condition of life for human hunter gatherers.

It follows that contemporary humans have the same developmental psychology as did hunter gatherers, so that if evolutionary theory can provide insights into the developmental psychology of hunter gatherers it will thereby provide insight into developmental psychology generally (see Cosmides and Tooby 1987 for an excellent discussion).

What will be important for our purposes will be the inferential patterns by which the motivation M is identified and the inferential patterns which support the hypothesis H_M that M is an innate and relatively non-malleable feature of human psychology. What I claim is that the identification of M is routinely achieved by a profoundly illegitimate inference in which (something very much like) the *evolutionary role* of a feature of human behavior is taken as the *propositional content* of a (usually unconscious) motive underlying it.

EXAMPLE: It is often suggested that altruism arose in early humans through *kin selection*: that although acts of altruism reduced the individual fitness of altruists, they increased the fitness of their kin sufficiently that genes which were expressed in altruistic behavior were favored. One way to describe this proposed evolutionary scenario is to say that humans "evolved a tendency to be altruistic towards their kin." Understood as a description of the evolutionary function of altruism among early humans, this is an accurate report of the scenario in question. It is, however, routine to fallaciously infer from scenario descriptions of this sort, that the (perhaps unconscious) motive which underwrites (most?) displays of human altruism is a concern for the altruist's kin (or members of his or her "in group"). For a presentation of this scenario, together with a different but equally fallacious inference to a motivational structure, see Barash 1979, p. 132-169.

EXAMPLE: In a widely cited paper, clearly from the era of "evolutionary psychology" rather than "sociobiology," Cosmides and Tooby 1987 come as close as anyone to making explicit the inferential pattern I am discussing here. They hold that "*The evolutionary function of the human brain is to process information in ways that lead to adaptive behavior*; the mind is a description of the operation of a brain that maps informational input onto behavioral output" (Cosmides and Tooby 1987, p. 282, emphasis theirs).

They elaborate this doctrine as follows.

> *When applied to behavior, natural selection theory is more closely allied with the cognitive level of explanation than any other level of proximal causation. This is because the cognitive level seeks to specify a psychological mechanism's function, and natural selection theory is a theory of function*. Natural selection theory specifies how an organism should respond to different kinds of information from its environment. It defines adaptive information processing problems that the organism must have some means of solving. Cognitive programs are solutions to information processing problems. (Cosmides and Tppby 1987, p. 285, emphasis theirs)

In explaining the application of this principle, they examine the implications of the kin selection scenario about altruism we have just discussed. They conclude as follows.

> ....an organism's behavior cannot fall within the bounds of the constraints imposed by the evolutionary process unless it is guided by cognitive processes that can solve certain information processing problems that are very specific. To confer benefits on kin in accordance with the constraints of kin selection theory, the organism must have cognitive programs that allow it to extract certain specific information from its environment: who are its relatives? which kin are close and which distant? what are the costs and benefits of an action to itself? to its kin?. The organism's behavior will be random with respect to the constraints of kin selection theory unless (1) it has some means of extracting information relevant to these questions from its environment, and (2) it has well-defined decision rules that use this information in ways that instantiate the theory's constraints. A cognitive system can generate adaptive behavior only if it can perform specific information processing tasks such as these. (Cosmides and Tooby 1987, p. 288)

Note that Cosmides and Tooby explicitly link the evolutionary function of adaptive behaviors to the *computational function* (and thus to the propositional content) of the underlying psychological states. What is important is that Cosmides and Tooby propose that one can infer the computational structure of the underlying psychology from the evolutionary function or role of the behavior, rather than merely that one can infer that the underlying computational structure must have been one capable of underwriting the predicted behaviors *in the EEA*. See section 2.4.0. for a discussion of the fallacy in inferences of the first sort.

The credibility of routine fallacious inferences of this form is facilitated, I claim, by the situation that many of the linguistic forms which are deployed in describing evolutionary scenarios or evolutionary hypotheses are either metaphorical uses of forms ordinarily employed to report propositional attitudes like desire, preference or purpose, or are ambiguous between propositional attitude expressing uses and their use in the making of evolutionary claims. The inferences I have in mind obtain their plausibility in part from the fact that they trade on the ambiguities involved.

EXAMPLE: Sometimes the evolutionary scenario just mentioned is expressed by statements like "Humans are altruistic *in order to benefit kin*." It is characteristic - indeed, I suggest, part of the meaning of such expressions as "in order to" in this literature - for authors of such claims to infer that the (unconscious) motive underwriting human altruism is a concern for kin. Wilson 1978 suggests that this inference would be appropriate, although in that book he dissents from the evolutionary scenario in question.

ANOTHER EXAMPLE: See the discussion of Daly and Wilson 1997 in the next section.

Similarly, some psychological/social terms ("altruism," "nepotism," and "strategy," as in "kin recognition strategy," for example) have been introduced *metaphorically* as technical terms in the study of the evolution of behavior. They retain, I shall argue, *as parts of their meanings in* *the literature in evolutionary psychology*, fallacious inference patterns appropriate to their literal, as opposed to their metaphorical, uses.

EXAMPLE: Phenotypic traits are said to be *altruistic* just in case they reduce the individual fitness of organisms that exhibit them while enhancing the fitness of their conspecifics. Some evolutionary psychologists reject the scenario for the evolution of altruism just mentioned in favor of a scenario according to which other-regarding or moral motivations were established by *individual* selection because they enhanced the effectiveness of those who had them in so far as participation in cooperative reciprocal arrangements was concerned (this is the "evolution of reciprocal altruism" *sensu* Trivers 1971). If some such scenario is correct, then altruism in humans was not *altruistic* in the metaphorical technical sense just discussed. It is routine (and, I suggest, part of the meaning of the term "altruism" in the literature) for defenders of this sort of evolutionary scenario to fallaciously conclude from the scenario in question that the *motivation* for apparently altruistic actions in humans is largely selfish.

In Wilson 1978, Wilson subscribes to this conception of the evolution of apparently altruistic motives. He concludes,on the basis of this scenario, that there probably is not much "hard core altruism." It is clear from the context that he intends this as a finding about *human psychology*. Nowhere in the book does he inform the reader that there is a technical (albeit metaphorical) use of the term "altruism" in evolutionary biology and that his arguments rely on using the term in this sense. Instead, he routinely grounds inferences about the presence or absence of genuine ("hard core," non-selfishly motivated) altruism in human psychology in considerations about the extent to which traits are altruistic in the metaphorical technical sense, without comment. This is precisely what one would expect if one thought that a conflation of the two meanings of "altruism" was part of the conceptual meaning of that term in the sociobiological literature.

Recall that the inferences we are discussing proceed from an evolutionary scenario involving a behavior B, to an alleged motivation M for B, and then to the hypothesis H_M that M is an innate and relatively non-malleable feature of human psychology. I have indicated some of the ways in which M is inferred illegitimately from the evolutionary function posited in the evolutionary scenario for B. It remains to explain how inferences to innateness and non-malleability are facilitated.

2.3. The Standard Pattern of Extrapolative Sociobiological Inference, II: Inferring the Innateness and Non-Malleability of Motives

Here the standard pattern of inference itself is less complex, but the roots of its credibility may be somewhat more complex. The pattern is simply that, once the motive M for B has been arrived at, it is inferred from the evolutionary scenario that M is innate and not very malleable.

EXAMPLE: In one famous and famously controversial example, the preference for kin or members of one's in-group which is posited as a result of a (fallacious) inference from the kin-selection scenario for human altruism is taken to be a manifestation of an evolved innate and relatively non-malleable xenophobia that, in turn, is offered as an explanation of the persistence of contemporary racism even in the face of reforms aimed at its elimination. This argument appears in Barash 1979 and is hinted at in Wilson 1975. Wilson 1978 rejects its conclusion because he adopts the individual selection scenario for apparent altruism discussed above - and concludes that we are motivationally largely selfish, but happily not innately xenophobic *as we would be if real ("hard core") altruism had been established by kin selection*.

EXAMPLE: Here is an example from the more recent literature, after "evolutionary psychologists" became less bold in their pronouncements about politically controversial matters than they were during the "sociobiology" period. In a widely cited study, Buss 1989 explores the predictions regarding human mate preference which, he maintains, can be made from evolutionary theory. He considers two different scenarios for the evolution of mate choice in human males, and draws different predictions from them about male mate choice *in general* (i.e., not just in the EEA). According to one scenario, early human males sought short-term mating partners. On this assumption, he expects that natural selection would have favored "a preference for females in their early 20s who show cues that are positively correlated with fertility(Buss1989, p. 177)."

According to the other scenario, early human males sought long-term mating partners. On that assumption, Buss takes evolutionary theory to predict a preference for "females in their mid-teens who show cues indicative of high reproductive value (Buss 1989, p. 177)."

It is plain from the text that Buss, without further argument, takes these predictions to be applicable to human males under conditions other that those of selection. For example, immediately after introducing the two hypotheses, Buss considers the views of other theorists. He describes the position of Williams 1975, without comment, as follows:

> Williams (1975), in contrast, predicts a compromise preference between reproductive value and fertility due to the existence of both long-term mating bonds *and* some possibility of *divorce* and extrapair matings. [First emphasis the author's; second emphasis mine.] (Buss 1989, p. 177)

Since *divorce* is, presumably, a particular social institution not always present in hunter-gatherer groups, it seems obvious that Buss understands the various evolutionary predictions he is considering to be understood as applying to conditions other than those which prevailed in the EEA.

That he understands the predictions of evolutionary theory in just that way is also made clear by his own experimental design. The data he brings to bear on the evolutionary predictions in question consist of the *contemporary* subjects' responses to questions about the parameters they consider important in mate choice. The data are cross cultural, and are obtained from subjects from 37 different *contemporary* cultures.

It is thus clear that Buss understands the various evolutionary scenarios to imply predictions about human motivational structures under a variety of conditions; his cross-cultural tests of the predictions are precisely designed to test for the presence of innate and largely non-malleable (rather than culturally specific) motivational structures. Thus his study reflects a paradigm case of the sort of inference from evolutionary scenarios to innate non-malleable psychological structures we are discussing.

It is worth noting that Buss nowhere defends (or even considers) this particular feature of his inferential practices. This is exactly what one would expect on the assumption defended here that inferences of the sort we are discussing are parts of the conceptual meaning of the relevant terms and can thus be presupposed, by authors and readers in the literature in question.

The plausibility of such inferences appears to have several sources. At least three different sources of plausibility seem to be involved, which deserve separate discussions.

2.3.0. Evolution, "Biological Bases," and Instinct: The "Two Natures" Inference One source is certainly the (culturally widespread) tendency to contrast "nature" with "nurture" and to identify the former with the "biological," and the latter with rationality and culture. According to this conception, behavioral flexibility is by and large to be identified with the effects of human rationality or human culture, whereas the biological aspects of human nature are identified with a motivational psychology of (relatively non-malleable)impulses or instincts. Roughly, this is the classical conception according to which human nature is composite from our "animal nature" and our "rational nature," with the former nature identified with the biological aspects of human psychology.

This sort of rationalization for inferences from evolutionary scenarios to non-malleable innate motivational structures is, I believe, at work whenever a biological story - evolutionary or not - about some feature of human behavior is summarized by saying that the behavior in question has a "biological basis." The inference that the behavior is rooted in an innate non-malleable motivational structure is essentially automatic - indeed such an inferential tendency is, if I am right, part of the meaning of the expression "biological basis" in the relevant research traditions.

EXAMPLE: The best illustration of this rationale at work may well be from the work of LeVay (1991, 1993), even though LeVay is not an evolutionary psychologist. LeVay 1991 reports differences in size of a particular brain structure (one of the cell groups in the interstitial nuclei of the anterior hypothalamus) in heterosexual and homosexual men and then adduces the biological basis thus uncovered for sexual orientation as providing *prima facie* evidence that sexual orientation in human males is largely unlearned and non-malleable. Wilson's (1978) claim that "biology" has "culture" on a short leash reflects the same dichotomy between the biologically based components of human psychology and behavior and the rationally or culturally determined ones.

2.3.1. Free Will, Biological Determinism, and Psychological Compulsion There is another related source of the plausibility of the inferences we are discussing which is harder to establish by reference to the literature, but which seems to me evident in conversations with students and others. I suggest that many times the inference from an evolutionary scenario regarding a behavior, B, to a conclusion about its non-malleable innate motivation is facilitated by a tacit pattern of fallacious reasoning along the following lines. B is an instance of *evolved* behavior; thus B is *biologically determined* and hence B is not *freely chosen*. Instead, B is the product of a *compulsion* or other difficult to resist urge, of which its author is perhaps unaware. It is this pattern of inference which sometimes emerges explicitly when someone says, for example, that people cannot be blamed (or blamed as much as one might think) for being territorial, or xenophobic or selfish, or whatever, because they *evolved* so as to have the trait in question.

2.3.2. "Evolved" (and thus) "Natural" "Inclinations" or "Tendencies" Another source of the plausibility of the inferences we are considering lies in systematic ambiguities in the uses of expressions like "evolved inclination," "natural tendency," and the like. In one use of this expression, a *natural tendency* or (somewhat less clearly) *natural inclination* is an *evolved* (and in that sense *natural*) dispositional property of an organism. Example: In some areas, where there is snow in winter but not during the rest of the year, some mammals have a *natural tendency* to change their coat colors with the seasons so as to be protectively colored.

In another perfectly standard use of the expression, a *natural tendency* (or *inclination*) is a feature of motivational psychology which is innate (or is likely to be learned under almost any conditions) and hard to extinguish. Example: People (perhaps) have a *natural tendency* to become angry when threatened.

When, according to an evolutionary scenario, a motivational trait *evolved* (and is thus *natural* in the first sense) it is often tempting (natural?) - indeed, I claim, it is part of the meaning of terms like "evolved" and "natural" in the relevant literature - to infer, fallaciously, that the trait is innate and non-malleable (*natural* in the second sense).

EXAMPLE: Here there is a famous example in the literature. In a widely cited paper, Daly and Wilson (1997) deploy the resources of evolutionary psychology to provide an explanation for an interesting fact about child neglect and child abuse *in contemporary industrial societies*. When couples marry, (or otherwise form a family) into which they bring their biological children from previous marriages or relationships, each partner is less likely to abuse or neglect his or her biological children than to abuse or neglect the children of his/her partner from the previous relationship. Daly and Wilson offer as an explanation our "usual inclination" towards such asymmetries which they say is predicted by optimality models in evolutionary theory. They do not spell out the evolutionary scenario in any greater detail, but it is evident that what optimality modeling might be taken to predict is a tendency or inclination (in the sense of a dispositional property) to favor one's own biological offspring *under conditions of the sort prevailing in the EEA*. What the explanation they offer for the relevant *contemporary* behaviors requires, on the other hand, is an innate and (relatively) non-malleable *motivation* to prefer one's own offspring - one which is exhibited under conditions quite different from those of the EEA. The fallacious inference from the posited evolved inclination to the relevant "usual" (apparently: innate and relatively non-malleable) motivational state is required for their paper to be intelligible. I emphasize that they do not make this inference explicit. They take it for granted, just as one would expect if such inferences were parts of the meanings of terms like "evolved" and "tendency" in the relevant discourse.

2.3.3. Behaviorism, Behavioral Ecology and the "Scientific" Study of Behavior It's worth remarking that this last source of plausibility for the fallacious inferences we are discussing - and perhaps the others as well - is enhanced in *its* credibility by the residual impact of behaviorism on research in extrapolative evolutionary psychology, and by the development of *behavioral* *ecology* as a research framework for the study of non-human animals.

Behaviorism - understood as a methodological approach to the study of behavior in which reference to mental or psychological processes is (allegedly) foregone in favor of purely behavioral descriptions - has had a much more persistent influence in studies of animal behavior than it has in the study of human psychology. In part this has been so because the adoption of behaviorist methods and behaviorist rhetoric has been central to the devices by which students of the behavior of non-human animals have sought to insulate their methodology from the tendency to anthropomorphize non-human animals. It has come to be seen as a mark of the *scientific* study of animal behavior that the researchers adopt a behaviorist perspective.

Of course, as philosophers and students of human cognitive psychology have recognized for several decades, it is not possible to make generalizations about the behavior of animals, human or non-human , without employing, at least tacitly, some conception of the source of those behaviors. One of the points which early critics of behaviorism as a methodology in human psychology emphasized was that when "behaviorist" psychologists deployed taxonomies of human behavior in proposing general theories, they were usually (and usually pretty transparently) importing into their methodology tacit assumptions about mental states and processes. Thus, in practice "behaviorist" human psychology did not represent the abandonment of theorizing about the mental. Instead, behaviorist psychologists, *believing falsely about their own methods that they were without presuppositions about the mental*, left largely unexamined (and uncriticized) the conceptions about mental states and processes which they tacitly adopted.

I suggest that much the same thing currently happens in extrapolative evolutionary psychology. Many of its practitioners are primarily students of (non-human) "animal behavior," or were trained in the behaviorist tradition associated with "animal behavior" studies. Although, as the examples cited above indicate, their methodology certainly involves the tacit positing of innate and relatively non-malleable (perhaps unconscious) motives, many of them (like earlier behaviorists) believe about their own practices that they do not involve positing mental states and processes. They are inclined to understand their "evolutionary" (or "optimality") predictions about human *behavioral dispositions* as being largely independent of psychological theorizing
Of course, the *correct* account of a phenotypic trait (whether psychological, neurological, or whatever) will be compatible with the *correct* account of its evolutionary history. Thus when evolutionary psychologists adopt an *explanatory* approach--when they seek to provide evolutionary explanations for aspects of human psychological phenotypes (like, e.g., innate and non-malleable motives) which have been *independently identified by psychological research*--they can reasonably expect that the results of their research will complement, rather than conflict with, the findings of psychological research. In the case of *extrapolative* human evolutionary psychology however the aim of research is to *do psychological research*: to identify general features of human psychology or behavior which persist under conditions very different from those in the EEA. Here there could not possibly be methodological independence from proximal psychology, since the inferences evolutionary psychologists make from evolutionary scenarios are proposed as a way of *doing proximal psychology*--of identifying enduring *proximal* dispositional or motivational features of human psychology, suitable for predicting or explaining behaviors outside the EEA. Nevertheless there is a widespread conviction among practitioners that evolutionary psychology and the study of proximal psychological mechanisms are methodologically independent..

In the context of such tacit but unrecognized psychological theorizing it is unsurprising that sometimes descriptions of the evolutionary function of behaviors, or "purely behavioral" characterizations of them, do double duty as descriptions of the propositional content of tacitly posited motivational states. It is equally unsurprising that the profound epistemic problems with the patterns of inference by which these states are (tacitly) posited should go unrecognized in the relevant literature. The deep structure of these inferences is largely invisible to most participants. [Note that Cosmides and Tooby 1987 are clear exceptions.]

There are, I believe, two other factors at work which enhance the effect of residual behaviorism in masking the problems with the inferences in question. In the first place, even the most behaviorist students of (non-human) animal behavior are aware that some (explicit or tacit) assumptions are always at work when researchers frame inductive inferences in terms of some particular way of taxonomizing behaviors. In the discipline of *behavioral ecology*, to which evolutionary psychology is very closely related (arguably as a sub-discipline) a conception of the appropriate taxonomy of behaviors has arisen which *is* largely independent of theorizing about underlying psychological mechanisms. According to that conception, behaviors are to be categorized in terms of what we might call *ecological* parameters: parameters which contribute to determining Darwinian fitness like, e.g., effects on food gathering efficiency, predator avoidance, access to mates, etc.. Evolutionary theory is then taken to predict that organisms will exhibit behaviors which are near optimal with respect to fitness. Thus, for example, it might be predicted that territorial defense behaviors would increase during the breeding season in those bird species in whose environments suitable nesting sites are uncommon.

Most behavioral ecologists and most evolutionary psychologists accept roughly this conception of the proper parameters for descriptions of behaviors. If (and only if) the broadly "adaptationist" optimality modeling approach to evolutionary theory which informs the literature is correct (an assumption I do not challenge here) then ecological parameters are the right ones *for predicting behaviors in the EEA*.

The apparent successes such ecological parameters in predicting the behaviors of non-human animals in the wild has apparently led many evolutionary psychologists to expect that human evolutionary psychology can be successfully carried out using the same sort of ecological parameters in characterizing human behaviors. In particular, they are inclined to believe not only that this would be possible in principle but also that current inferential practices in human evolutionary psychology generally conform to this standard: that psychological terms (like "altruism") used in evolutionary psychology are, in the final analysis, just metaphorical ways of describing behaviors and their impact(s) on parameters relevant to fitness, and not descriptions of posited mental states or processes.

We have just seen, in the examples cited above, that this is not so. On any plausible conception of mental states (in particular, on any functionalist conception) the inferences involved in those examples do posit (perhaps unconscious) motivational states.

Still, it is easy to see how researchers whose main training and/or research is in the study of non-human animals could misunderstand their own methods when they turn their attention to the human case. The use of ecological parameters in categorizing behaviors is appropriate in the study of non-human animals *precisely because* it is the main aim of researchers to characterize their behaviors in the wild - that is, in conditions which are, or duplicate, the EEA for their psychological/behavioral dispositions and capacities. It is a peculiar feature of the project of extrapolative evolutionary psychology that it aims to predict/understand the behaviors of humans in environments quite different from the EEA for human psychology. Many evolutionary psychologists fail to appreciate the fact that this new project places very different demands on the resources they use in taxonomizing behaviors, with the result that they make tacit assumptions about the underlying psychology of human behavior *while believing about their own practice that it conforms to behaviorist strictures*. Thus the *existence* of the inferential patterns we have been discussing has been largely (but not entirely; see Cosmides and Tooby 1987) unrecognized.

The other factor which, I believe, operates to enhance the influence of residual behaviorism is the peculiar ideological setting of reductionist approaches to the study of human behavior discussed in Section 4.3. It is part of a widespread conception of "objectivity" in such studies that consideration of psychological (as opposed to "biological"!) mechanisms in the study of human behavior represents a departure from scientific rigor and objectivity. That such a position is, in the final analysis, incoherent (since, according to the materialist approach presupposed in the literature, psychological mechanisms *are* biological mechanisms in the nervous system) does not preclude its having a profound methodological influence.

2.4. Reality Check

I have been arguing that there are fallacious inference patterns which are parts of the meanings of words and phrases like "evolved," "natural," "tendency," "altruism," "...in order to...," "biological basis," "strategy," and the like. Indeed, I propose to argue that these inference patterns are sufficiently fallacious, given the explicit findings of evolutionary theory and related disciplines, that the principles and inferential practices which are parts of the meanings of such terms present an essentially incoherent picture of evolved behaviors, rather than the approximately true picture posited by a benign conception of scientific meanings. Of course, my critique of the benign conception is fundamentally mistaken if the inferences in question aren't seriously fallacious. In case their fallaciousness isn't obvious, I'll argue for it here.

2.4.0. Behaviorism, Functionalism, and How Not to Infer Motives Let's consider first the inferences from evolutionary scenarios to underlying motives which proceed by deploying (something like) the relevant evolutionary function description as a description of the propositional content of the underlying motivational structure. As an example let's consider the inference which takes as a premise

(1) Early humans evolved to be altruistic towards their kin,

and reaches the conclusion that

(2) The (perhaps unconscious) motive for altruism was (or is - but we'll come to issues of innateness and non-malleability later) concern for one's kin.

We have seen that an (unsuccessful) attempt to follow the strictures of behaviorism and (thus) to speak only about behaviors rather than about their psychological bases can lead one to tacitly make inferences of this sort without recognizing that one has done so. As it happens, an appreciation of why behaviorism failed can help us to see why inferences like the one above are fallacious.

The least controversial aspect of the standard functionalist critique of behaviorism is that it is impossible to mirror, in behaviorist terms, the valuable explanatory richness which we get by offering explanations of behavior in terms of underlying psychological states because there is no simple correspondence between behaviors (described in purely behavioral terms) and underlying psychological states. The mapping from behaviors to motives is many-many.

In particular, it is central to the critique of behaviorism that, for any given pattern of behavior under a narrow range of stimulus conditions, there will be a large number of quite different scientifically reasonable explanations in terms of underlying motives, beliefs and other psychological states. That's why, even as first approximations, "operational definitions" for psychological states are inadequate.

Let's apply this most basic anti-behaviorist insight to the cases of inferences like that from (1) to (2). Call two psychological theories *behaviorally equivalent under conditions, C,* just in case they predict the same behaviors for humans under C. We can state the basis for the critique of behaviorism just discussed this way: for any narrow range of conditions, C, and any scientifically plausible general theory of human psychology, T, there will always be several scientifically plausible competitors to T which are behaviorally equivalent to T *under C* but which differ from T, and from each other, in the psychology they posit and the behaviors they predict under conditions *different from C*.

Now, suppose that we have accepted some evolutionary scenario concerning certain human behaviors in the EEA (that is: under the special conditions of life of hunter gatherer tribes). Let T be the hypothesis about the underlying psychology of those behaviors which is suggested by the sort of inference we are discussing. In so far as T is scientifically plausible, it will have several different scientifically plausible competitors which are behaviorally equivalent to it the under conditions which prevailed in the EEA, but which yield different predictions about human psychology and behavior under different conditions. Each of these theories will be equally compatible with the evolutionary scenario, so accepting that scenario provides no scientifically justified reason to prefer T over these competitors, and the inference to T is thus fallacious.

Thus, for example, premise (1) is compatible with the psychological theory affirmed in (2), but it is equally compatible with any plausible theory of the development of human altruism which, like it, would predict that *in early tribal societies* altruistic behaviors would mainly benefit the kin of altruists. Given the plausible view that small tribal groups were mainly groups of kin, a plausible theory behaviorally equivalent to T might not have to posit any special psychological mechanism to explain the kin bias in early human altruistic behaviors. An alternative which did posit such a mechanism could posit any plausible mechanism such that - under the conditions of early tribal life - the factors which conduced to displays of altruism towards a person were *correlated with* kinship, like, e.g., familiarity.

This result is quite general (as the reader is invited to check for herself): for any evolutionary scenario of the sort popular among evolutionary psychologists there will be lots of importantly different scientifically plausible psychological theories behaviorally equivalent under conditions prevailing in the EEA to the one posited as a result of the standard pattern of inference from evolutionary functions descriptions to descriptions of the propositional content of motives. Thus such inferences are scientifically unjustified.

2.4.1. Dualism, Free Will, and the Scope of Tendencies: How not to Identify Instincts. We have just seen that the standard pattern of sociobiological inference from evolutionary scenarios to the propositional contents of motivational states is fallacious. It remains to examine the equally standard pattern of inferring that the motivational states (however identified) which underlie the behaviors specified in such scenarios are innate and non-malleable. Let's again focus on the sample inference involving altruism and kin-regard, this time with the commitment to innate and non-malleable "instinct" made explicit:

(1) Early humans evolved to be altruistic towards their kin,

and reaches the conclusion

(2) The (perhaps unconscious) motive for altruism is (non-malleably) a concern for one's kin or for members of one's in-group.

Of course, the anti-behaviorist functionalist considerations just rehearsed show that this inference is fallacious. In particular they show *both* that (even if the evolutionary scenario is right) the motivation for altruism in early humans need not have been a concern for their kin (or for their in-groups) *and* that, even if it were, that motivation need not have been innate or non-malleable. What I want to emphasize here is that the various considerations, rehearsed above, which might initially be thought to lend plausibility to the inference in question do not in fact provide any rational support for it.

Consider first the rationale which identifies evolved motives as *biologically based* and then infers that they are instinctual rather than being subject to rational or cultural influence. From a suitable dualist perspective - one in which our "animal natures" are aspects of the physical world whereas our "rational natures" are not - this inference might be justified.

As it is, however, from the tacitly materialist perspective regarding human psychology which (properly, I believe) underwrites any evolutionary study of human psychology, the opposite conclusion follows. Evolved features of our psychology have, of course, a biological (and thus a material) basis. Indeed *all* features of our psychology have a biological basis, since our mental life is a feature of the activity of our nervous systems.

It follows, however, that the dualist picture which assigns instinctual motives to the biological features of our psychology and rationally or culturally malleable motives to our non-material minds *can't* be right: we *have no* non-material mental life. Thus from the fact that a feature of our motivation "has a biological basis" nothing whatsoever follows about the extent to which it is innate or culturally or rationally malleable. The innate and non-malleable and the learned and malleable (and the innate and malleable and the learned and non-malleable, for that matter) are on an ontological par: they all have a "biological basis."

Similar consideration show that the rationale which rests on considerations about free will makes a fundamental mistake about the relationship between ontological and psychological questions. Suppose, for the sake of argument, that incompatiblism is true, so that the physical determination of behaviors makes them unfree in whatever sense is relevant to the question of free will. It follows from the assumption that the motivational basis for a pattern of behavior evolved that the relevant motivational structure is biological, and thus physical. So, assuming incompatiblism, it follows that the relevant behaviors are unfree (surely whatever level of quantum uncertainty there is in biological systems will not restore free will if free will is really compromised by physical determinism).

So, it follows that evolved behavioral patterns are unfree. What does not follow is *anything* about the psychology of those patterns. Incompatiblism, which we have assumed for the sake of argument, is *not* the view that physically determined actions are always the product of instinctual desires or compulsions, nor the view that physically determined action patterns are not subject to rational deliberation or cultural influence. Instead, it is the view that - whether they are innate or learned, rationally calculated or instinctual, culturally malleable or not - physically determined actions are unfree is some metaphysically important sense. Physical determination which proceeds *via* rational deliberation or *via* the causal influence of culture is, according to incompatiblism, just as much a source of unfreedom as determination *via* the operation of innate and non-malleable motivational structures.

Thus a thinker with incompatiblist views who infers that behaviors with a "biological basis" must derive from an innate and non-malleable motivational structure because they are unfree engages in fallacious reasoning.

Finally, let us turn to the inferential rationale involving appeals to evolved behavioral tendencies. Consider again the inference involved in Daly and Wilson 1997 from an evolutionary scenario involving special concern for one's own offspring to the conclusion that humans have a "usual inclination" to care preferentially for their own biological offspring. From the evolutionary scenario Daly and Wilson must have in mind (they do not state it explicitly) one can rationally infer that early humans had an *inclination* (in the sense of a dispositional property) which, *in the EEA*, resulted in their caring for children who were their own biological offspring, and did not (often) result in their caring for other children. This is the only "usual inclination" which one can infer from the evolutionary scenario: in fact the assertion of the existence of this "inclination" simply restates the scenario in question.

The point just made - that only the most limited sort of "inclination" or tendency can warrantedly be inferred from the evolutionary scenario in question - follows, of course, from the functionalist and anti-behaviorist critique presented earlier. All that the evolutionary scenario allows us to infer is that the psychology of early humans was such that, *coming to maturity in the EEA* they had some adult psychology or other which led *in the EEA* to the predicted pattern of child care. Nothing about innateness or non-malleability can be inferred, nor can it even be inferred that a special concern for one's biological children *because they are one's biological children* characterized the adult psychology of early humans.

It is important to see how conflating the notion of an inclination as a dispositional property exhibited under specific conditions (like conditions of the EEA) with the notion of a usual or natural inclination in the psychological sense can lead one to make the fallacious inference which Daly and Wilson make. In the *psychological* sense of "inclination," to attribute to humans a usual or natural inclination to do something *is* to attribute to humans a motivational state which could reasonably be expected to be exhibited under a wide range of conditions; that's what "usual inclination" means. So, if one mistakenly believes that the evolutionary scenario Daly and Wilson have in mind predicts an *inclination*, in that sense, to care preferentially for one's own offspring, then one will mistakenly conclude that the scenario leads to the prediction they advance.

2.4.2. Incoherence I intend to argue that the inference patterns we have been discussing are parts of the meanings of the key terms in evolutionary psychology and that, as such, they are counterexamples to the initially plausible idea that meanings in mature scientific disciplines are generally *benign*: that such meanings tend to reflect genuine insights into the relevant subject matter of the sort which would ordinarily be expected to contribute to the establishment of methodological commensurability between different research traditions with the same subject matter. It is thus important for me to argue that the inferences in question are very far from reflecting insights into the evolution of human psychology. In fact, what I propose to show is that these inference patterns, when taken together with basic evolutionary principles which are also candidates for components of meanings of the relevant terms if any things are, give rise to a conception of human psychology and its evolution which is essentially incoherent. It is this last claim which I propose to defend in the present section.

We have already seen that the inference patterns in question are seriously fallacious. Might it still be the case that the picture of the evolution of human psychology prevailing in evolutionary psychology is basically insightful, despite the prevalence of these inference patterns?

That the answer is "no" is indicated by two different considerations. In the first place, the inference patterns we have been discussing are not peripheral to evolutionary psychology. In fact, they are what defines the discipline and distinguishes it from other approaches to human psychology. Secondly, if made explicit, the inference patterns in question are profoundly inconsistent with the most basic components of evolutionary theory - with, that is, conceptions which are presupposed in any scientific discipline in which evolutionary theory is applied.

Consider first the question of the centrality of the inferences in question to evolutionary psychology (or "sociobiology"). It is important to see that what is distinctive about evolutionary psychology as a disciplinary project or research strategy is not simply the idea that human psychology is the product of evolution. That presupposition is fundamental, for example, to comparative psychological studies generally and to comparative neuropsychology in particular. It is a broadly evolutionary conception of the origins of human and non-human psychological structures which justifies the assumption that comparative studies will be informative with respect to the human case.

What IS distinctive of evolutionary psychology is precisely the research strategy of deploying the resources of optimality models and associated evolutionary scenarios as sources of insight into human psychological structures, especially motivational structures. But, the inferential patterns we have been discussing are precisely the methodological tools by which the alleged insights are obtained. In so far as they are seriously fallacious the central methodological practices of the enterprise of evolutionary psychology is seriously compromised.

In that regard, then, the conception of human psychology and its evolution reflected in these aspects of (what I claim are) the meanings of key terms in evolutionary psychology are anything but insightful. Worse, there just isn't any even remotely coherent (much less correct) conception of such matters reflected in the meanings of such terms. I assume here that the most fundamental explicit principles of evolutionary theory are components of the conceptual meanings of the relevant terms: they are certainly presupposed in the literature in such a way that someone who had not engaged with them would find it unintelligible.

So, the question of the coherence of the conception of the evolution of human psychology reflected in the meanings of the key terms in the literature of evolutionary psychology amounts to the question of whether or not the fallacious inferences we have been discussing could be part of a coherent (even if false) conception which also included the most basic principles of evolutionary theory. The answer is "no."

To see this, first consider the inferences from evolutionary scenarios which proceed by confusing the evolutionary function description of a posited pattern of behavior with a description of the propositional content of the motivational state which causes it. What such inferences require is that when expressions which are metaphorical uses of psychological descriptions (like "...in order to..." or "altruism") or terms ambiguous between psychological and non-psychological uses (like "inclination") are used to describe the adaptation of behaviors to the environments in which they were displayed one takes those expressions as expressions of genuine motives. Now one of the things which is explained in every basic course in evolutionary biology is precisely that Darwin showed that it is *never* appropriate to interpret the purposive or teleological language which we find it so natural to use in describing adaptations as literally describing motivational or purposive or teleological structures.

Thus the fallaciousness of the inference rules we are considering follows trivially from absolutely central principles of evolutionary theory. There is no coherent story which incorporates the appropriateness of those inference rules and the central principles in question.

The same is true of the inferences from evolutionary scenarios to the conclusion that the underlying motivations for the relevant behaviors are innate and non-malleable. One of the points explicitly acknowledged by evolutionary psychologists even if they tacitly ignore it (see e.g., *Lumsden* and Wilson 1981 and Cosmides and Tooby 1987) is also central to any conception of the evolution of learning capacities in humans or any other animals. It is that natural selection can operate to favor *learned* patterns of behavior. If this were not so, there would be no way in which natural selection could favor the establishment of learning mechanisms themselves.

Let me explain. Of course, it is true that when selection favors learned behaviors (for example, behaviors which depend on learned motives) it must also favor innate learning mechanisms whose operation underwrite the relevant learning, but neither the behaviors themselves nor the motives for them need be innate or non-malleable. In fact, if, when natural selection favored behavioral patterns those patterns were always unlearned and grounded in innate and non-malleable psychological states, there could not be any evolutionary explanation for the evolution of the capacities involved in learning itself. There would be no evolutionary function for learning mechanisms to play.

This point follows easily from the most basic principles of evolutionary theory and is (as I mentioned) explicitly recognized even by evolutionary psychologists who routinely ignore it in their inferential practices. Thus there is no coherent conception (much less an insightful one) which incorporates this second feature of the standard pattern of evolutionary psychological inference and acknowledges the most basic principles of evolutionary biology. If the inferential principles in question and the most basic evolutionary principles are parts of the meanings of the key terms in evolutionary psychology, then we have an actual and important case in which meaning of terms in a mature scientific research tradition are anything but benign.

2.4.3. Meanings? It remains to examine the question of whether or not the inferential practices we have been discussing - and the basic principles of evolutionary theory - are constituents of the meanings of the key terms in evolutionary psychology. This is, of course, an empirical question about the sorts of conceptual engagements which are required for someone to understand the relevant literature and to appreciate methodological and theoretical discussions in evolutionary psychology. I'll present evidence for the claim that these components of the conceptual resources of evolutionary biology are constituents of meaning in the required sense but, of course, the reader may choose to conduct her own investigations into this question.

The key reason for thinking that the inferential principles we are discussing - as well as the most basic principles of evolutionary theory - are constituents of meaning is that, in the literature in question and in professional discussions in evolutionary psychology they are (a) tacitly presupposed and (b) central to the argumentative strategies of the papers and discussions in question. Since they are rarely made explicit (which is hardly surprising given the conflict with fundamental tenets of evolutionary theory) the inferential practices in question must be presupposed as appropriate (in the case of uncritical engagement) or explicitly identified by the critical reader/listener as central to the relevant literature (in the case of critical engagement) in order for the reasoning engaged in by professional evolutionary psychologists to be intelligible. This is the key mark of components of conceptual meanings.

Other factors also indicate the appropriateness of thinking of the inferential practices in question as constituents of meaning. If engagement with a particular conceptual resource within a research tradition is centrally important to the intelligibility of that tradition, and if the standard mode of engagement within the tradition is uncritical engagement, then one would expect that the typical participant in the tradition would have difficulty understanding claims which denied the cogency of the conceptual resources in question.

I now report - as a summary of 20 years of experience teaching a course on methodological issues in sociobiology/evolutionary psychology - that this is indeed the case. Especially for students who have been exposed to this research tradition in biology courses it is extremely difficult to even remember the broad outlines of the sorts of criticisms of the inferences in question which I have presented here.

I want to emphasize that I am not reporting merely that such students fail to be convinced by those criticisms. That would be unsurprising in a context in which teachers they regard as *prima facie* authoritative disagree about fundamental methodological matters. What I am reporting is that students who have been made familiar with the literature in question have very great difficulty even paraphrasing the criticisms in question.

Similar results obtain, I now report, when philosophical critics of evolutionary psychology encounter able practitioners of that discipline in methodological discussions. the response of those practitioners is precisely what one would expect if the inferential practices in dispute were, for the evolutionary psychologists, components of the very meanings of the relevant terms/concepts. In my experience, evolutionary psychologists respond to criticisms of the sort we have been discussing by acknowledging the logical possibility that the conclusions of the relevant inferences might be false, but by expressing incredulity that the critic would take this fact to underwrite a serious methodological criticism. That is, they respond to the criticisms in question just as one should if responding to someone who rejected a fundamental principle of inductive inference on the grounds that it was not deductively valid. What they - like their students - are unable to do is to appreciate the fact that the criticisms in question are not exercises in philosophical skepticism but ordinary scientific critiques. That, I suggest, is true precisely because, for practitioners who are uncritically engaged with the inferential practices in question, those practices are components of the meanings of the key disciplinary terms and thus define the limits of relevant scientific criticism.

Here's another way to think about the same question. No one who works in evolutionary psychology would report her view that some motivational feature of human psychology was established by individual selection by saying that it was not *altruistic* unless she intended to convey to her listeners/readers that the motivation for that feature was not (or was not mainly) other regarding, *even though* the term "altruism" has, in the literature in question, the technical meaning "contributes to the reproductive fitness of conspecifics while reducing the fitness of the organism that exhibit it." The inferential pattern which involves (in effect) conflating this technical sense of "altruism" with the standard psychological sense is so much a standard of reasoning within the discipline that it is an essential component of communicative practices.

Still, someone might say that the inference in question was not part of the *meaning* of the technical term "altruism" but rather that the fact about communicative practices I have just mentioned is simply a matter of conversational *implicature*. Here's the rebuttal. This "implicature" has the property that it takes a semester long graduate seminar to cancel it, and the cancellation will not be successful with respect to some members of the seminar. They will not be able to remember from day to day that there are two senses of "altruism" involved. Similar results obtain for the other standard inferential patterns in evolutionary psychology. If one is interested in having a theory of communication, "implicatures" which are that hard to cancel *are* constituents of meaning.

2.4.4. Malignancy? We need also to address the question of just how far from benign the meanings of the key terms in evolutionary psychology are. Recall that the benign meaning thesis is not the thesis that the meanings of scientific terms will ordinarily be exactly true (if they are doctrines) or perfectly reliable (in the case of methods) but only that the meanings of scientific terms in mature sciences can be expected to reflect sufficient insight into the relevant subject matter that, when two traditions share a common subject matter, the meanings of the relevant terms in the two traditions can be expected, in general, to contribute to - rather than to detract from - the establishment of methodological commensurability between them.

This is pretty plainly not the contribution which the meanings of the key terms in evolutionary psychology make to the prospects of commensurability between that research tradition and those (like traditional anthropology, sociology, political economy and social psychology) which address many of the same issues about human social behavior. The key inferential patterns which are among the relevant meanings in evolutionary psychology make it extremely difficult for its practitioners to appreciate the plausible methodological criticism arising from competing traditions as anything other than (a) a desperate retreat into skepticism in the face of the new science of evolutionary psychology or (b) a reflection of a failure to appreciate the implications of evolutionary biology. The consequence is that the prospects for fruitful dialogue upon which the prospects for commensurability depends are greatly reduced. Meanings in evolutionary psychology are genuinely malignant.

3. LESSONS, I: MEANINGS, REFERENCE, AND SCRIPTS

3.0. Meanings

We now need to extract some philosophical lessons from our examination of meanings of theoretical terms in sociobiology/evolutionary psychology. In the first place, we have seen that scientific terms do sometimes have conceptual meanings. There may be domains of discourse or research traditions about which the Quinean dictum is true that there is no real distinction to be drawn between those principles or inferential patterns which are so central to practice that they count as components of meaning and those whose centrality is insufficient for them to qualify for that honor. Nevertheless this is by no means always true. In the case of sociobiology/evolutionary psychology, for example, there *are* inferential practices such that engagement with them is essential to any understanding of the most basic inferences and arguments in the literature and these *are* components of the conceptual meanings of the terms and phrases involved. There is a place in the semantic theory for scientific terms for a notion of conceptual meaning.

Second, although it may often happen that doctrines or inferential patterns become established as components of conceptual meaning because they reflect important insights into the relevant subject matter(s), this is not generally true even for well established scientific research traditions. *The* central inferential practices which are components of meaning in sociobiology/evolutionary psychology are deeply fallacious.

Indeed, the components of the meanings of the key terms in a scientific research tradition need not provide anything like a coherent conception of its subject matter: the meanings of key terms in sociobiology/evolutionary psychology, for example, do not.

For this reason, the meanings of terms in a scientific research tradition can be barriers to the establishment of methodological commensurability with respect to other traditions with the same subject matter, just as Kuhn suggested. The meanings of scientific terms, even in a well established research tradition, can be *malignant* rather than *benign*.

3.1. Meanings and Reference

Once it is acknowledged that scientific terms sometimes have meanings it is natural to suppose that they figure in the establishment of reference for such terms in approximately the way suggested by traditional descriptivist conceptions of reference: that a scientific term will refer to whatever property, magnitude, or whatever its meanings are *approximately* true of (in the case of doctrines which are components of meaning) or approximately reliable about (in the case of inferential practices). At least it is plausible to think that a constraint of this sort is part of what determines the reference of a scientific term, with the rest of the work being done by relevant causal relations between instances of the referent in question and use of the word in question.

Descriptivism isn't even that close to being a correct account of reference. As the case of sociobiology/evolutionary psychology shows, the meanings of a scientific term within a research tradition need not be such that they could be approximately true of, or insightful about, *anything*. They can provide a fundamentally incoherent picture of the subject matter in question.

It is true, of course, that reference is an epistemic success phenomenon (Boyd 1989, 1993, 1999a). In all but the most contrived cases a term, t, refers to an entity, e, only if inferential (and perceptual) practices in the relevant community are such that there is some explanatorily important tendency for what is predicted of t to be true of e's. What the examples from sociobiology/evolutionary psychology indicate are two points about the epistemic success which is thus central to reference. First, the practices whose epistemic reliability underwrite reference need not be the practices engagement with which is central to the possibility of successful communication: they need not include the components of the conceptual meanings of the terms in questions.

Secondly, even when scientific terms are used in mature scientific research traditions with all the trappings of "paradigms" in Kuhn's sense - graduate programs, refereed journals, professional societies, and (most importantly) shared standards of evidence - it need not be the case that the epistemic successes which underwrite the referential connections between the vocabulary of the "paradigm" and features of the world are mainly achieved by practices within the paradigm. The term "altruism," as it is used within sociobiology/evolutionary psychology refers (some of the time) to a real other-regarding aspect of human motivation. This can only be so because the use of the term "altruism" is to some extent determined by inferential practices which are epistemically reliable with respect to that other-regarding motivational state.

As it happens, sadly, the methods of sociobiology/evolutionary psychology are not among those practices: they are very poor methods for finding out about altruism. So, in effect, evolutionary psychologists who use the term "altruism" to refer to an other regarding motivational state are *borrowing* the reference of this term from more reliable practices which are situated in other psychological research traditions and in everyday "folk psychological" practice.

3.2 Reference and Progress

Part of the import of Kuhn 1970 was to undermine the Whig conception that the operation of scientific methodology can, almost always, be expected to lead to progress, so that, when scientists fail to make such progress the appropriate diagnosis of their situation is, at least ordinarily, that they failed to apply scientific methods correctly.

Part of the Kuhnian challenge to that conception is available just as a consequence of the recognition of the theory-dependence of scientific methods: whether or not progress is made with respect to a family of scientific questions depends, not (just) on whether or not practitioners conscientiously practice scientific methods, but on whether or not the background theories and conceptions which determine their methods happen to be relevantly approximately true. Since theory-independent inductive methods do not exist, the expectation of progress at any point in the history of any scientific discipline must rest on the profoundly *a posteriori* estimate that the background theories in question are relevantly approximately true. So the expectation of progress in science needs to be significantly qualified.

There is, of course, an additional dimension to Kuhn's critique of the Whig conception. According to Kuhn there is no progress - in the sense of successively closer approximations to the truth - during "scientific revolutions." Kuhn's reasons for denying that there is this sort of progress have - as we have seen - a semantic component: the key terms in scientific paradigms are supposed to change meanings *and referents* during "revolutions," so that it *can't be* that there is progress towards a more accurate understanding to the common subject matter of the earlier and latter paradigms *because they don't have a common subject matter*!

Naturalistic rebuttals to Kuhn's conception of reference for scientific terms refute this last claim and establish that terms in quite different research traditions can share a common subject matter, so that *Kuhn's* semantic argument against the expectation of progress across revolutions fails.

Since reference is a matter of some degree of epistemic success, it might be reasonable to conclude that when the key terms in a scientific research tradition refer there is reason to expect that the research methods within that tradition will be epistemically reliable enough to underwrite a *prima facie* expectation of progress.

What we have just seen is that, perhaps surprisingly, a semantic argument closely related to those offered by Kuhn shows that this expectation is not generally speaking justified. In at least some important cases the methods, and the doctrines which determine the methods, within a research tradition - including those whose centrality is such that they properly count as components of the meanings of the key terms - are so misleading that they cannot be expected to underwrite progress (or progress towards commensurability with other traditions sharing the same subject matter). In such cases there will, of course, be epistemic successes upon which the reference relations for the various key terms supervene, but they need not lie within the purview of the tradition in question: they can be almost entirely borrowed.

Thus the causal or naturalistic theory of reference for scientific terms underwrites a general expectation of scientific progress only in the most highly qualified sense: scientific terms used within a research tradition have determinate referents only if *somewhere or other* perceptual or inferential practices which govern the uses of those terms produce *some sort of* approximations to the truth. But, these practices need not lie within the domain of the research tradition in question, *or of any scientific research tradition at all*. It is fully compatible with a causal/naturalistic conception of reference for the theoretical terms of some scientific disciplines or research traditions that most of the relevant epistemic successes which underwrite reference are achieved outside of scientific practice. Whether or not things are ever this bad, the expectations of scientific progress which follow from just causal theories of reference are very limited indeed.

3.3. Quinean Scruples Revisited

We may restate some of the points just made in a way which illuminates the ways in which Quinean criticisms of the notion of meaning fail to apply to the questions about conceptual meanings which we are addressing here. According to the traditional empiricist conception of meanings which Quine challenges in rejecting the analytic synthetic distinction the meaning of a general term (like a theoretical term in science) is provided by a set of conceptual entities - statements, criteria of application, or inferential principles - which:

a. are analytic or analytically grounded,

b. are thus rationally unrevisable,

c. determine the most basic rational norms for inferences involving the terms in question, and

d. are such that acceptance of them is definitive of linguistic/conceptual/communicative competence with respect to the terms in question.

What the Quinean critique of the analytic synthetic distinction suggests is that there are (almost) no terms, and certainly no theoretical terms in science, which are associated with conceptual entities satisfying a.-d. because (almost) no terms of any sort are associated with conceptual entities satisfying a. and b.

We may think of a.-c. as specifying that the sorts of meanings anticipated by the empiricist tradition are supposed to be *methodologically normative* in an especially strong sense. They are supposed to be (analytically) methodologically immune from revision and to (analytically) set the most basic inferential standards for the terms in question.

d., by contrast, requires of meanings only that they be *communicatively normative*: that they underwrite the conditions for mutual intelligibility of statements and arguments.

Now Quine's arguments that there are (at least for almost all words) no conceptual entities satisfying a.-c. do not address the question of whether or not there are such entities satisfying d., alone. All they establish is that whatever *communicatively normative* conceptual entities there might be associated with a word (in the sense of satisfying d. with respect to it) they will never (or almost never) satisfy the standards which a.-c. set for *methodological normativity* (because nothing will).

Similarly, when Quine argues against the prospect that a clear distinction can be drawn between conceptual entities central enough to the use of a term to count as parts of its meaning and those conceptual entities which are merely well established, his arguments are entirely directed towards questions of methodological normativity: he argues essentially that there is a continuum of levels of immunity from revision for the conceptual entities associated with a word at a time, but that, because no conceptual entities rise to the analytic level of immunity, there is no non-arbitrary level of immunity along this continuum for a given term above which such entities are methodologically normative enough to be counted as components of meaning.

I suppose that this is entirely right, but it does not address the central question of the present essay regarding communicatively normative meanings. What I have been arguing is that for lots of theoretical terms in science there *are* conceptual entities which are unproblematically components of their conceptual meanings - which are communicatively normative for those terms. I have not argued that there is a sharp boundary between such entities and those which are not components of conceptual meaning, but only that there are (at least in some cases) clear cut cases of such components.

3.4. "Conceptual Roles" and Two Kinds of Normativity

The points just made depend on the fact that, once the analytic-synthetic distinction is abandoned for scientific terms, there are two distinct sorts of normativity which can be attributed to conceptual entities. On the one hand, they may be said to be normative in a methodological sense, such that someone who affirms that they are normative in that sense is making a methodological judgment regarding the (approximate) truth, or reliability of the conceptual entity in question. To say of a doctrine that it is normative in this sense is to say that it reflects some important truth or approximate truth about the relevant subject matter; to say that an inferential pattern is methodologically normative is to affirm its reliability and justifiability with respect that subject matter.

By contrast, to describe some conceptual entity as communicatively normative is *not* (as it would be on the standard empiricist conception of the semantics of scientific terms) to endorse it in this way,but merely to indicate that engagement with it (uncritical *or critical*) is centrally important to one's ability to understand the argumentation and reasoning in the scientific tradition. As the treatment of the standard inferential practices in sociobiology/evolutionary psychology in the present essay indicates, it is possible to argue that conceptual entities are communicatively normative while insisting that they are very, very, far from being methodologically normative.

The distinction between the two sorts of normativity is related to corresponding distinctions between conceptions of the *conceptual role* of a scientific term and between conceptions of the *rationality of inferences* in science. Conceptual role first. The conception of conceptual meaning I have offered here is a species of the genus of conceptual role accounts of term meanings. In the first place, I have insisted that certain inferential roles (which are the paradigm cases of "conceptual roles") are components of the meanings of scientific terms. Secondly, in identifying doctrines as components of the meanings of such terms I have emphasized the question of whether or not engagement with them is crucial to an understanding of the central inferences and arguments within the relevant research tradition (again a question of conceptual role).

Where the proposal offered here differs from that reflected in more traditional empiricist conceptual role semantic theories is that, for those theories, the conceptual roles which are parts of the meanings of scientific terms are analytically justified, so that to identify a conceptual entity as a component of the conceptual role of a scientific term is to provide for it the strongest possible justification. On this sort of conception, conceptual role meanings are as benign as conceptual entities could possibly be.

By contrast, on the conception proposed here the conceptual roles in terms of which the meanings of scientific terms are defined are those which would be identified by the true theory of communicative practices within the relevant research tradition rather than by the (analytic components of) the true theory of the relevant subject matter. So one need not be endorsing a conceptual role in identifying it as a central component of the meaning of a term. Scientific terms can - and do - often have conceptual meanings that they shouldn't have.

It is not part of my thesis here that we cannot have a different sense of conceptual role according to which the conceptual role of a theoretical term is the conceptual role which it ought to have. My only claim is that, if we pay attention to those the insights of traditional conceptual role semantics which survive the critique of the analytic synthetic distinction, we will see that conceptual roles in this latter sense play no significant role in theories of communication or commensurability in science.

All of the distinctions we have examined in the present section rest on the rejection of the analytic-synthetic distinction for theoretical terms. The rejection of that distinction is closely related (just how closely we need not figure out for present purposes) to the rejection of the conception that there are *a priori* justifiable, and thus theory-neutral methods in the sciences. As we have seen it is the fact that there are no such methods which makes it possible in the first place that methodological incommensurability should obtain between competing scientific research traditions or "paradigms," even when their practitioners are scientifically rational.

Recognizing that the methods of science are theory-dependent in this way permits us to draw one more important distinction. Recall that I maintained that the components of the conceptual meanings of theoretical terms are *communicatively* but not *methodologically* normative: in identifying a conceptual entity as a component of the conceptual meaning of a theoretical term one need not *endorse* the methodological or theoretical role it plays.

There is, however, a sense in which the components of the conceptual meanings of theoretical terms must be methodologically normative: no conceptual entity associated with a term in a research community is going to be such that engagement with it is crucial for understanding inferences and arguments unless *by the standards prevailing in that research community* the entity in question plays a very important methodologically normative role.
Once we see this, we can see that it is a reflection of an important phenomenon of the *relativity of scientific rationality* to the doctrines and practices prevailing within a research tradition at a time. The relationship between scientific rationality and epistemic success is complex. Roughly, a scientist or scientific community acts rationally just in so far as they conscientiously apply the methods dictated by (what their careful and thoughtful judgment takes to be) the best available theories. Practices which are rational in this sense are epistemically reliable *only* to the extent that these best available theories are relevantly approximately true.

So, scientific rationality is (in the sense specified) relative to a research tradition. There need be no failure of individual or collective scientific rationality when rational scientific practices (in unfortunately situated research traditions) fail to reliably resolve scientific issues or (as in the sorts of cases we are considering here) fail to establish commensurability between competing research traditions.

Thus, for investigations of the relationship between the semantics of scientific language and of the phenomenon of incommensurability have uncovered two respects in which Kuhn was right: there are sometimes barriers to methodological commensurability arising from differences in the meanings of the same terms in different research traditions *and* this fact is crucially related to the phenomenon of tradition-relativity of scientific rationality. Note, however, that nowhere have we had any occasion to invoke a relativity of truth or of "reality" or of the "worlds" scientists study.

3.5. Insulated Scripts and Incoherent Meanings

I have argued that the conceptual meanings of key terms in sociobiology/evolutionary psychology are essentially incoherent. A natural question arises of how it is possible that a system of inferential practices and substantive doctrines which, taken together, provide no coherent picture of the relevant phenomena whatsoever could be sustained as methodological norms in a serious scientific discipline practiced by many intelligent and sophisticated people. Why isn't the incoherence fairly readily detectable - and thus correctable?

Part of the answer, of course, is that the inferential practices are not made explicit. Roughly, the relevant "rule" of inference is this: From a premise which attributes a particular evolutionary function, F, to a class of human behaviors in the EEA, infer that an innate non-malleable motive to accomplish F is part of human psychology and underwrites analogous behaviors in conditions other than those prevailing in the EEA *unless* this conclusion is implausible in the light of what is already known about human learning mechanisms. In the latter case, infer that there is an innate and non-malleable psychological mechanism at work and posit for it a structure as closely related to an innate and non-malleable motive to do F as you can without sounding silly.

Now no one would subscribe to this inferential principle if it were spelled out explicitly. It is not (but see the discussion of Cosmides and Tooby 1987 in section 2.2.). Instead, I suggest, what happens is that students who are acquiring the inferential standards prevailing in sociobiology/evolutionary psychology become familiar with, and learn to deploy, what we might call *inferential scripts* reflecting the standard applications of the standard sociobiological/ evolutionary psychological inferential strategy. They acquire, through practice and through reading the literature, a "feel" for the sorts of extensions of those scripts which are treated as legitimate innovations within the research tradition. This tacit knowledge of the standards prevailing in the tradition constitutes their (uncritical) engagement with the inferential practices in question.

This is, I believe, a special case of a phenomenon recognized by Kuhn which helps to underwrite his choice of the term "paradigm." According to Kuhn assimilation into the community of researchers whose work constitutes a paradigm crucially involves mastering certain key "exemplars" of appropriate research and acquiring (largely) tacit understanding of the prevailing standards for innovations within the framework represented by those exemplars. In the case of the standard inferential practices in sociobiology/evolutionary psychology what is involved is tacit knowledge of the prevailing standards for the acceptability of new variations on the received applications of the standard inferential practice.

One further point will help us to understand how a basically incoherent conception can set the inferential standards for a scientific research tradition. In my (perhaps caricatured) account of what the explicit form of the standard inferential practice would be I characterized the general inferential pattern and then added a clause to the "rule" which instructs the inferer to depart from the standard practice when necessary in order to avoid scientific embarrassment. Without this "loophole" the rule would dictate inferences which would be profoundly - and obviously - fallacious in the light of cases in which there is selection favoring learned behaviors, and thus selection for learning mechanisms connected with motivations and behaviors rather than for innate motives and behaviors.

In practice, of course, practitioners do not apply an explicit (and methodologically peculiar) rule which contains a clause recommending avoiding embarrassment. Instead, the tacit understanding of the methodological standards prevailing within the research tradition which they have acquired includes a tacit understanding of when the more general standard inferential strategy is to be invoked and when it is not. This tacit understanding *insulates* the general inferential strategy from the embarrassment which would result from its consistent application. This phenomenon of insulation of inferential practices helps to explain why the incoherence of the conceptual resources of the research tradition remains invisible to practitioners.

The exact psychology and sociology of the establishment of tacit inference principles and of their insulation are empirical issues which go beyond the scope of this paper. For philosophers however, some appreciation of how the relevant mechanisms work can probably be gleaned from an appreciation of the role of skeptical arguments in philosophy, especially within the empiricist tradition. It is routine for philosophers to deploy skeptical arguments against metaphysical or normative doctrines to which they are unsympathetic, even when those arguments are so powerful that, *if they were consistently applied*, they would result in the rejection of knowledge claims to which the philosophers in question are committed (for a discussion of this sort of selective skepticism in empiricist philosophy of science see Boyd 1983, 1985a; for a discussion of the same phenomenon in metaethics, see Sturgeon 1984).

In such cases, the philosophers have acquired - through their training in philosophy or their understanding of the relevant literature - a tacit understanding of the skeptical inferential practices in question: an understanding which embody a suitable (tacit) insulation of those practices from potentially embarrassing applications. Something like the same sort of tacit understanding operates to establish insulation in cases of incoherent meanings in scientific research traditions.

4. LESSONS, II: DISCIPLINARY BOUNDARIES, IDEOLOGY AND THE PROSPECTS FOR COMMENSURABILITY

4.0. How Widespread is the Phenomenon?

We have seen that scientific terms have conceptual meanings, and that their meanings can be malignant with respect to commensurability, in much the way Kuhn thought, even though his arguments for referential incommensurability fail. We've also learned that a certain plausible "descriptivist" proposal within a broadly naturalistic conception of reference - that when scientific terms have conceptual meanings they can be expected to be approximately true/reliable - is subject to important counterexamples.

An important question obviously remains: how widespread is the phenomenon of malignant meaning? This is, of course, an empirical question whose resolution would require investigations that go beyond the scope of the present paper. What I propose to do here is to identify some features of the situation of sociobiology/evolutionary psychology which I think may plausibly have contributed to the emergence of malignant meanings, and to suggest that in addressing the question of how widespread malignancy of meaning is it will be important to examine other research traditions or other frameworks of inquiry which share these features.

I suggest that among the key factors responsible for the emergence of meaning malignancy in the case of sociobiology/evolutionary psychology are three interrelated phenomena: the *disciplinary isolation* of specialized research traditions like sociobiology, the *ideological setting* within which its research takes place, and the (ideological) *prestige of reductive science*. I'll explore these factors in the following sections and suggest, on the basis of my exploration, that the prospects for the eventual achievement of commensurability in the face of malignant meanings may, in some important cases, depend on social and political factors external to the institutional practices of science.

4.1. (Sub)Disciplinary Isolation

One of the ways in which it is possible to see that the standard inferential patterns in sociobiology/evolutionary psychology are fallacious is to recognize that they rest on tacit premises and methods regarding human (and non-human) psychology which have been rejected for good reasons in the critiques of behaviorism which underwrote the development of contemporary "cognitive science" approaches to human and non-human psychology. I have suggested that the plausibility of many (perhaps all) of the fallacious inferential practices in question rests in part on the influence which those behaviorist premises and methods still have in programs in "animal behavior" within which much of the research in sociobiology/evolutionary psychology is conducted.

What is important for our purposes here is that, roughly speaking, it is possible to become a professional evolutionary psychologist without becoming familiar enough with the literature in the more "cognitive science" literature to anticipate or appreciate the tensions between one's own (sub)disciplinary inferential strategies and those of researchers with a keener appreciation of the problems with behaviorism. Arguably, at least, the way in which this (sub)disciplinary isolation insulates the inferential practices in evolutionary psychology from anti-behaviorist criticisms helps to explain how the fallacious inferential practices in question could become sufficiently normative *within sociobiology/evolutionary psychology* that they emerged as components of the conceptual meanings of its key terms.

In so far as this is true one might expect to find (some of) the preconditions for malignant meanings whenever subdisciplines develop their own journals, graduate programs, professional organizations and the like largely independently of such institutions within the relevant broader disciplinary context. At least arguably, a tendency towards this sort of subdisciplinary isolation is a consequence of very widespread tendencies towards specialization, and of the institutional arrangements which grow out of competition between subdisciplinary approaches for funding and prestige. In this regard, then, the prospects for malignant meanings seem unfortunately favorable.

4.2. The Opposite of Isolation: Ideological Embedding

I am inclined to think, however, that the primary contributor to the emergence of malignant meanings in sociobiology/evolutionary psychology lies in a way in which inferences in that research tradition are *not* isolated from the conceptual elements of other approaches to the same subject matter. What I have in mind is the ways in which the standard inferential pattern in sociobiology/evolutionary psychology is ratified by prevailing ideology in the culture at large.

I mean to speak here of *ideology* in both senses of the term: in the sense in which it refers to conceptual resources which are central to a particular cultural setting at a time *and* in the sense in which it refers to those aspects of ideology in the first sense whose prevalence is explained by their role in the ratification of existing patterns of power and wealth.

With respect to ideology in the first of these senses, it is important to note that the general conceptions (or misconceptions) of rational and animal natures, of behaviors with a "biological basis," of "nature" and "nurture," of "free-will," etc. which partly explain the plausibility of the standard inferential strategy in sociobiology/evolutionary psychology are paradigm examples of prevailing ideology in the first of these senses, so much so, in fact, that some of the fallacious inferences we have been discussing might plausibly be said to be constitutive of the meanings of the terms mentioned above in everyday intellectual discourse. That many of the fallacious inferences in the research tradition are thus underwritten by elements of the broader culture surely serves to enhance their immunity from the criticisms that one might otherwise expect to be forthcoming within the discipline itself.

What seems especially important in the case of sociobiology/evolutionary psychology is that both its inferential strategies and its findings are embedded as well in an ideological setting of the second, more distinctly political sort. It is a characteristic feature of even much of recent (and more "moderate") work in evolutionary psychology that its inferential principles and its conclusions underwrite and are underwritten by the sort of cynical conception of human nature and of human potential which have been the premises and the conclusions of social Darwinism since Darwin. They thus carry with them the special credibility which, as the history of science indicates, attaches in any modern historical period to the scientific rationalizations of (the least agreeable and least justifiable of) its social structures and relations.

The combination of these two aspects of ideological embedding constitute, I conjecture, the primary explanation for the fact that an essentially incoherent combination of doctrines and inferential practices has come to constitute the conceptual meanings of the key terms of sociobiology/evolutionary psychology. There is, however, an additional ideological factor which, I conjecture, is also important.

4.3. The Prestige of Reductive Science

Another factor which tends to insulate the inferential practices of sociobiology/evolutionary psychology from the scrutiny which might undermine their credibility - and which thus helps to explain how they maintain their status as components of conceptual meaning - must surely be the epistemic and methodological prestige which currently attaches to reductive explanatory approaches in the sciences. My own experience teaching about methodological issues in evolutionary psychology indicates that the very fact that the approach in this research tradition involves a "reductionist" research strategy serves to lend its methods and findings a credibility quite independently of the details of the arguments presented in the relevant literature.

This, too, is a bit of prevailing ideology. Its roots are, of course complex. It seems evident that, to some extent, the spectacular success of biochemical approaches to issues in genetics, physiology, neurochemistry and the like have contributed to this prestige. So has the greater prestige of the "hard" sciences (of which biology is now one) in comparison with the "soft" sciences from which alternatives to sociobiological/evolutionary psychological explanations are likely to arise. At least arguably, these factors are ideological mainly in the first of the two senses we are discussing. [I actually think that they are ideological in the second sense as well, and importantly so, but it is beyond the scope of this essay to develop this point.]

What I want to emphasize here (because my experiences in teaching this material and in arguing with practitioners of evolutionary psychology suggests to me that it is very important) is an aspect of the prestige of reductionist explanations in the human sciences which is plainly ideological in the political sense. To a remarkable extent there is an association between (1) the conception of an objective scientific opinion on matters human, (2) a reductionist approach to such matters, and (3) "toughmindedness" regarding them, where the latter carries simultaneously the connotation of unbiased investigation and the implication that only a cynical conception of such matters can possess the appropriate level of objectivity.

The connection between (2) and (3) is partially explained by the fact that reductionist explanations for unfortunate features of human social practice *will* characteristically result in a cynical conception of human nature, because what characterizes reductionist explanations of that sort is that they attribute the unfortunate features of human social relations to properties of individual human psychology rather than to, e.g., features of social or economic organizations. But the general association of objectivity with reductionism and with toughminded cynicism is a much deeper feature of prevailing social ideology. It reflects a gendered (masculine) conception of scientific objectivity and underwrites cynicism about the prospects for social progress in general.

This association of the authority of (alleged) scientific objectivity with the defense of prevailing relations of power and wealth is, of course, the paradigm case of the political ideological role of appeals to scientific authority in modern societies. It would be unwise to underestimate the extent to which this association undermines the prospects for the sorts of criticisms of the (cynical) standard inferential strategy in sociobiology/evolutionary psychology which would be required to overcome the malignant meanings and make possible methodological commensurability with other approaches to human nature.

The examination of these ideological effects suggests that malignant meanings might be more likely to be found in cases in which (sub)disciplinary isolation is associated with ideological embedding, especially of the political sort. If this guess is correct, then such (sub)disciplines as psychometrics, applied game theory and the psychology of individual differences might be places to look.

4.4. The Prospects for Commensurability: The Politics of Criticism

The question I want to address now is what the prospects are for eventually achieving commensurability between competing research traditions when the conceptual meanings of their shared terms are malignant in one or more of them. Recall that, as I understand the term, *commensurability* obtains between two research traditions only if there are methodological resources which are fair to both traditions *and epistemically reliable*, and which are adequate to adjudicate the differences between them. For example, in the case of the establishment of commensurability between sociobiology/evolutionary psychology and other approaches to human psychology it would be necessary for practitioners of the former research tradition to come to appreciate the deep problems with those standard inferential practices which are currently components of the conceptual meanings of their key theoretical terms.

There is surely no informative generalization which applies in all cases of malignant meanings, but I think that an important lesson can be drawn regarding cases in which there is important ideological embedding underwriting the sustenance of malignant meanings. In the case of sociobiology, as it was called at first, there was a period (roughly from the publication of Wilson 1975 until shortly after the publication of Kitcher 1985) during which a dialogue of sorts emerged in which practitioners of sociobiology and their critics (especially Lewontin 1976 and Gould and Lewontin 1979) engaged in exchanges which were accessible to substantial parts of the intellectual community and which had the effect of producing a *quite temporary* pause in the growth of the plausibility which sociobiological theorizing enjoyed within the scientific community generally.

What has happened subsequently has been the re-establishment of the credibility of evolutionary psychology (as it is now called), *with the standard (fallacious) inferential practices largely unchanged*. The effect of the criticisms by Gould, Lewontin, Kitcher and others was *not* to undermine those inferential practices, even though the critics showed (along lines roughly like those explored here) that those practices are deeply fallacious and indicated their ideological roots. Instead, the only significant and lasting effect of these (quite cogent) criticisms was to reduce the frequency of sociobiological (I mean "evolutionary psychological") speculations about racism and war and to replace these "hot" topics with reproductive strategies, sex differences, cooperation, and altruism. To a limited extent this took some of the political "edge" off the research traditions, but it left the inferential practices - and their ideological roots - largely untouched.

No doubt the story of why the important criticisms raised during the "sociobiology" period had such limited effects is quite complicated, but one factor seems definitely to have been important. The criticisms in question emerged during a period of radical critiques of political ideology in intellectual life. The very possibility of the articulation of those criticisms - and of their intelligibility in the intellectual community generally - rested on political developments (the emergence of militant anti-racist, anti-imperialist and anti-sexist movements) largely external to institutional science.

Had these movements not emerged, I suggest, the ideological embedding of the standard sociobiological inference practices would have been sufficient to make challenges to them hard to invent and even harder to present persuasively. The return to almost unchallenged status (and thus to the role of components of conceptual meanings) of those practices is largely explained, I suggest, by an equally external phenomenon: the worldwide defeat of leftist movements of the sort which helped to underwrite the critiques in question. Whatever the explanation may be for that defeat, it was certainly not largely a matter of considerations internal to institutional science or to rational application of scientific methodology.

What this suggests, and what I now propose, is that in cases in which malignant meanings are sustained by embedding in social ideology, the prospects for the establishment of commensurability may depend (much) more on the emergence of political movements embodying critiques of social ideology than on any developments internal to the research traditions in question.

4.5. Politics and the Relativism of Rationality

I should emphasize two ways in which the point just made about the relationship between social ideology and malignant meanings illustrates important dimensions of the relativity of scientific rationality. In the first place, consider the situations of researchers who are pursuing a malignantly flawed research program as a result of ideologically sustained commitments to mistaken doctrines or fallacious methods. In the absence of a change in *political* circumstances which reduces the conceptual and cognitive impact of the relevant bits of social ideology, even the exercise of perfect (theory-dependent) scientific rationality need not suffice to permit them to recognize the flaws in their methods or doctrines *even if* there exist scientific arguments against these doctrines and methods which (a) are actually presented somewhere in the literature and (b) would be persuasive were it not for the ideologically sustained malignant errors in question. The question of the possibility of error correction and progress under such circumstances is, in an important sense, a political question.

This so because there is *always* an important political dimension to the reliability or unreliability of scientific methods. Because they are ineliminably theory-dependent, scientific methods are systematically reliable only when, and to the extent that, there are available background theories which are themselves relevantly approximately true. But, this is not a sufficient condition for their reliability. Appropriate social, economic and political arrangements must exist for relevantly accurate theories to be disseminated and for work properly grounded in them to be supported and its results assigned an appropriate level of credibility (see Section 5.1).

This, in turn, happens within a research tradition - under ordinary circumstances - *only* when powerful economic and social interests are served by the promulgation of approximately true answers to the relevant questions. When powerful interests would be ill served by the truth, the various mechanisms of social ideology operate in such a way that - under ordinary conditions - the products of institutional science serve those interests at the expense of the truth.

This pattern is obscured when we direct our attention only to the evident reliability of the methods of the sciences in those areas of the physical sciences and biology where the most important achievements of contemporary science have emerged. If we fail to recognize that these are precisely the areas with respect to which there are industrial and military interests which are served by the discovery of (approximate) truths, we can fail to recognize the role of economic and political interests in determining the institutional and ideological structures in which these successes are made possible.

Examining cases like that of sociobiology/evolutionary psychology, in which those very factors conduced to the systematic scientific ratification of ideology rather than to uncovering the truth, helps us to see that the prospects, such as they are, for progress towards the truth depend crucially on political factors in both sorts of cases.

The lesson of the present discussion is that there are important cases in which the operation of scientific methodology - even when practiced conscientiously and rationally by smart people - is inadequate, in the absence of political transformation, to overcome the ways in which social ideology introduces malignancy into findings, methods, and even concepts and meanings in the sciences.

5. BROADER IMPLICATIONS

5.0. Commensurability and Incommensurability in Ethics

Standard relativist arguments against moral realism are similar in structure to Kuhn's arguments for referential incommensurability: they identify conceptual differences between different traditions of moral inquiry and practice, treat these as representative of differences in meaning, and conclude that the traditions in question cannot share a common subject matter. I have elsewhere (Boyd 1988) maintained that these arguments are not decisive against moral realism and that an appreciation of the semantics and epistemology of scientific inquiry can help us to see that moral realism is a serious possibility in metaethics.

What I want to do here is to indicate ways in which - even if moral realism is correct (as I believe it is) - there may be insights as well as errors in the relativist treatment of metaethical issues. If we think of moral reasoning not primarily as the province of professional philosophers but as something engaged in by members of various different social, religious, national, and professional groups then, like sociobiology, moral inquiry within many of these groups will have the property that those who engage in it are largely insulated from critiques from inquirers in other groups. Moral inquiry certainly shares with inquiry in the human sciences the property that it is embedded (deeply embedded!) in social ideology.

Thus moral inquiry shares the properties which, I conjecture, are especially likely to give rise to malignant meanings and consequent incommensurability. I want to explore briefly some further consequences of these properties of moral inquiry.

First, it is important to see that the embedding of moral inquiry within political ideology is manifest along several different dimensions. One is obvious: there is a systematic tendency for the prevalent moral views at a time to be such that their acceptance serves the needs of the powerful. Indeed, so strong is this tendency that the debunking critique of morality attributed to Marx has much to recommend it (see Boyd 1988, 1995; Miller 1984; Wood 1972, 1979, 1984; Gilbert 1981, 1982, 1984; for discussions).

Secondly, although morality need not (I think should not) be thought of as having a religious foundation, there is in fact an intimate connection between moral inquiry and religious institutions and practices. In so far as these are instruments of political ideology, their ideological embedding will extend to moral inquiry as well.

Moreover, whatever else moral inquiry may be about, it certainly involves inquiry about human nature and about the effects of possible social policies on human well being; so it is about human social structures as well as about human nature. Thus moral inquiry centrally includes topics which are, along with religion, the paradigm *loci* of the influence of political ideology.

Finally, consider the relative isolation of communities of moral inquirers from each other which, I am suggesting, can be expected to contribute to the emergence of malignant meanings and incommensurability. This isolation is itself partly a phenomenon of ideology in the political sense. A central social function (*the* social function, if you're inclined to be pessimistic) of moral discourse and practices is to protect the interests of the powerful. [I think that Marxists are right to think of these interests as the interests of ruling classes, but people with radically different conceptions about how political and economic power work can agree that morality serves the powerful.]

The capacity of morality to serve this ideological function is, in turn, greatly enhanced by the phenomenon of isolation of communities of moral inquirers. Consider, for example, cases in which racist anger is directed at the allegedly immoral behaviors or character traits of isolated minorities in order to deflect criticisms from unpopular economic policies, or in which chauvinistic moral condemnation of the "enemy" is used to rationalize unjust wars. Neither sort of ideological deployment of moral sentiments would serve its function of social control with nearly the same effectiveness in situations in which there was an appreciation, on the part of members of the target group for the propaganda, of the moral perspective and outlook of the scapegoated group or of the "enemy." It is thus unsurprising that ordinarily social mechanisms are in place to create and rationalize the required isolation between moral communities. Such isolation is a component of the (political) ideological role of appeals to morality.

I am thus inclined to think that the battery of ideological forces arrayed against the prospects for progress in moral understanding, and thus against the establishment of commensurability between moral communities, is much more substantial even than those arrayed against progress and commensurability in many of the human sciences (not, perhaps more substantial than in: military and diplomatic history, the economics of poverty, the psychology of gender differences, or the genetics of intelligence). It follows that, even on the moral realist assumption (which I endorse) that it has a real subject matter, moral inquiry may be especially vulnerable to the phenomenon of incommensurability and that overcoming incommensurability may crucially involve political struggle against those interests whose power prevailing ideology rationalizes.

5.1. Concluding Radical Postscript

In a certain sense the question of commensurability and incommensurability with respect to competing paradigms or research traditions is the question of the possibility of *objectively* resolving the differences between them. In the present essay I have followed standard scientific realist practice in arguing that Kuhn's arguments for referential incommensurability fail to demonstrate permanent incommensurability for the cases he considers and I have offered a conception of the possibility of the emergence of commensurability as a result of suitable dialectical interactions between practitioners in competing paradigms between which there are not initially the methodological resources for incommensurability.

Nevertheless I have made a gesture towards Kuhnian arguments, and towards the sorts of postmodernist relativism they are often taken to support, by agreeing with Kuhn, and with many postmodern thinkers that:

1. Scientific (and other) terms have conceptual meaning,
2. there can be conflicts of meaning between competing paradigms or research traditions, such that

3. they seriously compromise the near term prospects for commensurability.

I have argued as well that

4. The phenomenon of malignant meanings and of the incommensurability they can occasion are indicative of a sort of relativism about scientific rationality.

I have also agreed with a point more often made by postmodern thinkers than by Kuhnian ones, namely, that

5. Features of social ideology can be parts of the conceptual meanings of scientific (or other) terms.

I want to explore the relationship between the points made in the present essay and the proper understanding of the notion of objectivity. It is part of the stereotype of objectivity that objective methods are theory-neural and thus immune to social ideology (and to individual idiosyncracy as well). I have suggested, in my discussion of the scientific prestige of reductive explanations, that there is a complex ideological association between scientific objectivity, reductionist explanations and "toughminded" cynical conceptions in the human sciences. These features of the prevailing conception of objectivity are, of course, matters of political ideology: the associations with reductionist toughmindedness are part of a broader pattern of scientific rationalization for the prevailing relations of power and wealth, and the conception that science is "objective," and *thus immune from ideological influence* itself serves the *ideological* function of enhancing the credibility of the political rationalizations produced by the sciences.

There is, however, such a phenomenon as scientific objectivity and - just as our stereotypical conception suggests - it serves to explain those cases in which scientific methodology is systematically and often spectacularly epistemically reliable. Scientific objectivity is a matter of the reliable operation of scientific methods to systematically produce real knowledge, and scientific methods *do* exhibit this property under a variety of historically important circumstances. Where our stereotypical conception is wrong is that it gets the identification of those circumstances essentially backwards.

Scientific methods, when they are reliable, are reliable because they are theory-dependent *and* because the background theories upon which their applications depend are relevantly approximately true, *but* even the availability of relevantly accurate background theories will not render the methods of science reliable *unless* the scientific disciplinary practices in question are situated appropriately with respect to sources of political and economic power.

In ideological contexts like the present one, in which scientific findings serve to ratify profound inequities in power relations, ordinary institutional science can function in an epistemically reliable way *only* when political developments largely external to institutional science create temporary shifts in the balance of political influences within scientific institutions. It is not an accident, for example, that critiques of biological determinist cynicism about the prospects for social and political progress - in so far as they have any significant impact on mainstream professional life in the human sciences - do so only in the context of large scale and (largely) external political struggles for the progress in question. This was, for example, the case with respect to critiques of sociobiological methods and with respect to related critiques of the professional literature on the genetics of intelligence (see, e.g., Block and Dworkin 1976).

In so far as scientific objectivity regarding politically controversial issues in the human sciences is possible in the absence of such large scale movements, it is always within the context of oppositional movements which substitute a politically less misleading context for that which obtains in mainstream scientific institutions. The role which scholars close to the Communist Party played in the U.S. in the 1930's and 1940's in keeping alive a tradition of objective study of racism and its effects provides a useful example here.

Thus, far from depending on insulation from politics, objectivity about politically controversial matters in the human sciences (and elsewhere) is, when it obtains, *always* a political achievement. Consider now cases in which powerful interests sustain scientific institutions in which methods reliably lead to real knowledge, when that sort of epistemic success is in their interests. The establishment and maintenance of such institutions, and of their epistemic reliability (when it is wanted), is *also* a political achievement of sorts, even in those cases - military research for example - where epistemic successes are not especially to be admired. Scientific objectivity, when it obtains, is thus *always* a (partly) political achievement.

So, scientific objectivity is a theory-dependent, politics-and-power dependent phenomenon, and the view that it arises under conditions of theory-neutrality and political immunity is itself a piece of ideology in the political sense of the term.

Now many postmodernist critics of science have maintained that the notion of scientific objectivity is an essentially political or ideological notion, and that it is a mistake for critics of the *status quo* or of the ideologies which support it to deploy the notion of scientific objectivity. They often hold that the notion of objectivity (and, for example, the notions of truth and of knowledge) are political tools of oppression, inappropriate for use in the context of ideology critique.

Obviously I dissent from this view, but I think that the resources of the present essay allow us to appreciate an important grain of truth in it. I am inclined to think that the ideological conception of objectivity we have been discussing is, in contemporary educated circles (at least in the U.S. - I haven't sampled adequately elsewhere), constitutive of the conceptual meaning of the term "objectivity." This is, I believe, one of those cases in which ideologically determined elements are pretty clearly parts of the meaning in academic discourse of a politically important term.

If this is so then it is true, as critics of objectivity maintain, that the *concept* of objectivity, as that concept is manifest in contemporary intellectual discourse, is a tool of oppressors, at least in so far as its applications in politics and in the human sciences and related areas of inquiry are concerned.

On the other hand, the *phenomenon* of objectivity - the epistemically reliable deployment of scientific methods - is essential to projects of social criticism and ideology critique. In the human sciences and related domains, scientific objectivity cannot - absent wholesale political and economic change - be achieved within the normal working of institutional science, except in rare contexts of sustained political struggle. It can be obtained outside the normal workings of institutional science only in the contexts provided by oppositional movements.

It is probably also true that only in the contexts provided by political struggle can the concept of objectivity be freed from its malignant meanings. In almost every respect the critics of scientific objectivity are thus correct about the depth of the ideological association between the concept of scientific objectivity and processes of economic and political oppression. They are right as well to think that the situation can only be remedied politically. But they need not - and indeed must not, if they are to succeed - abandon the projects of criticism to which concepts of objectivity, knowledge, and truth are essential.

In an era in which we are routinely called upon to celebrate the defeat of the left and the emergence of a globalized capitalist economy, it is a good idea to reflect that the points just made about the special epistemic role of political struggle and of oppositional political institutions are part of the legacy of Marxist (and in particular of Leninist) political theory. They have not outlived their usefulness.

Cornell University

NOTES

REFERENCES

Alexander, R. (1979). *Darwinism and Human Affairs*. Seattle: University of Washington Press.

Alexander, R. (1987). *The Biology of Moral Systems*. New York: Aldyne de Gruyter.

Allen, Elizabeth, et. al (1978). "Against 'Sociobiology'," in Caplan, ed. *The Sociobiology Debate*, pp. 259-268. New York: Harper and Row 1978.

Angier, Natalie. (1978). "Study of Sex Orientation Doesn't Neatly Fit Mold," *New York Times July 18, 1993.*

Barash, D. (1979). *The Whisperings Within*. New York: Harper and Row.

Betzig, L., ed. (1997). *Human Nature: A Critical Reader*. New York: Oxford University Press.

Bleier, Ruth. (1986). "Sex Differences Research: Science or Belief?" from Bleier, ed. *Feminist Approaches to Science*, pp. 147-164. New York: Pergamon.

Block, Ned J, and Dworkin, Gerald, eds. (1976). *The IQ Controversy*. New York: Pantheon.

Boyd, R. (1982). "Scientific Realism and Naturalistic Epistemology." in P.D. Asquith and R.N. Giere (eds.) *PSA 1980. Volume Two*. E. Lansing: Philosophy of Science Association.

Boyd, R. (1983). "On the Current Status of the Issue of Scientific Realism." *Erkenntnis* **19**: 45-90.

Boyd, R. (1988). "How to be a Moral Realist." in G. Sayre McCord (ed)., *Moral Realism*. Ithaca: Cornell University Press.

Boyd, R. (1985a). "Lex Orendi est Lex Credendi." in Churchland and Hooker (eds.) *Images of Science: Scientific Realism Versus Constructive Empiricism*. Chicago: University of Chicago Press.

Boyd, R. (1985b). "Observations, Explanatory Power, and Simplicity." In P. Achinstein and O. Hannaway (eds.) *Observation,Experiment, and Hypothesis In Modern Physical Science*. Cambridge: MIT Press.

Boyd, R. (1989). "What Realism Implies and What It Does Not" *Dialectica*.

Boyd, R. (1990a). "Realism, Approximate Truth and Philosophical Method" in Wade Savage, ed. *Scientific Theories, Minnesota Studies in the Philosophy of Science*. vol. 14. Minneapolis: University of Minnesota Press

Boyd, R. (1990b). "Realism, Conventionality, and 'Realism About'" in Boolos, ed. *Meaning and Method*. Cambridge: Cambridge University Press.

Boyd, R. (1991). "Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds." *Philosophical Studies* **61**: 127-148.

Boyd, R. (1992). "Constructivism, Realism, and Philosophical Method." in John Earman, ed. *Inference , Explanation and Other Philosophical Frustrations*. Berkeley: University of California Press.

Boyd, R. (1993). "Metaphor and Theory Change" (second version) in A. Ortony (ed.) *Metaphor and Thought*. 2nd Edition. New York: Cambridge University Press.

Boyd, R. (1995). "Postscript" to "How to be a Moral Realist," in P. K. Moser and J. D. Trout, eds. *Contemporary Materialism: A Reader*. New York: Routledge.

Boyd, R. (1999a). "Kinds as the "Workmanship of Men": Realism, Constructivism, and Natural Kinds." *Proceedings of the Third International Congress, Gesellschaft für Analytische*

Philosophie. Berlin: de Gruyter.
Boyd, R. (1999b). "Homeostasis, Species, and Higher Taxa," in R. Wilson, ed. *Species: New Interdisciplinary Essays*. Cambridge: MIT Press.
Brink, D. (1984). "Moral Realism and the Skeptical Arguments from Disagreement and Queerness." *Australasian Journal of Philosophy*. **62.2**: 111-125.
Brink, D. (1989). *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
Buss, D. M. (1989). "Sex Differences in Human Mate Preferences: Evolutionary Hypotheses Tested in 37 Cultures," *Behavior and Brain Sciences* **12**, 1-14, as reprinted in Betzig 1997 .
Carnap, R. (1950). "Empiricism, Semantics and Ontology." *Revue international de philosophie*, 4th year.
Cosmides, Leda and Tooby, John. (1987). "From Evolution to Behavior: Evolutionary Psychology as the Missing Link," in John Dupre, ed. *The Latest on the Best: Essays on Evolution and Optimality*, pp. 277-306. Cambridge: MIT Press.
Daly, M. and M.Wilson. (1997). "Child Abuse and Other Risks of not Living with Both Parents," in L. Betzig ed. 1997. *Human Nature: A Critical Reader*. New York: Oxford University Press.
Feigl,H. (1956). "Some Major Issues and Developments in the Philosophy of Science of Logical Empiricism." in H. Feigl and M. Scriven (eds.) *Minnesota Studies in the Philosophy of Science*, vol. 1. Minneapolis: University of Minnesota Press.
Field, H. (1973). "Theory Change and the Indeterminacy of Reference." *Journal of Philosophy* **70**: 462-481.
Gilbert, A. (1981). "Historical Theory and the Structure of Moral Argument in Marx," *Political Theory* **9**: 173-205.
Gilbert, A. (1982). "An Ambiguity in Marx's and Engel's Account of Justice and Equality," *American Political Science Review* **76**: 328-46.
Gilbert, A. (1984). "Marx's Moral Realism: Eudaimonism and Moral Progress," in J. Farr and T. Ball, eds. *After Marx*. Cambridge: Cambridge University Press.
Goodman, N. (1973). *Fact Fiction and Forecast*, 3rd edition. Indianapolis and New York: Bobbs-Merrill.
Gould, S. J. and R. C. Lewontin (1979). "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Program." *Proc. Roy. Soc. Lon. (B),* **205**: 581-598.
Gould, Stephen Jay. (1978). "E.O. Wilson, On Human Nature," *Human Nature*, volume 1, number 10 Dec 1978: 20-28.
Gould, Stephen Jay (1981). *The Mismeasure of Man*. New York: Norton.
Hanson, N.R. (1958). *Patterns of Discovery*. Cambridge: Cambridge University Press.
Hoyningen-Huene, P. (1993). *Reconstructing Scientific Revolutions: Thomas S. Kuhn's Philosophy of Science*. Chicago: University of Chicago Press.
Kitcher, Philip. (1985). *Vaulting Ambition: Sociobiology and the Quest for Human Nature*. Cambridge, Mass.: MIT Press.
Kripke, S.A. (1971). "Identity and Necessity." in M.K. Munitz (ed.) *Identity and Individuation*. New York: New York University Press.
Kripke, S.A. (1972). "Naming and Necessity." in D. Davidson and G. Harman (eds.) *The*

Semantics of Natural Language. Dordrecht: D. Reidel.
Kuhn, T. (1970). *The Structure of Scientific Revolutions*, 2nd edition. Chicago: University of Chicago Press.
LeVay, Simon, (1993). Untitled Short Editorial Piece, *The Nation*, July 5, 1993.
LeVay, Simon (1991). "A Difference in Hypothalamic Structure Between Heterosexual and Homosexual Men," *Science* vol. 253 (30 August 1991): 1034-1037.
Lewontin, Richard C. (1976). "Sociobiology--a Caricature of Darwinism," *PSA 1976*, vol 2: 22-31.
Lumsden, C. and E.O. Wilson. (1981). *Genes Mind and Culture: The Coevolutionary Process*. Cambridge, Mass.: Harvard University Press.
Pinker, S. (1996). *How the Mind Works*. Cambridge: MIT 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.
Putnam, H. (1972). "Explanation and Reference." in G. Pearce and P. Maynard, eds. Dordrecht: Reidel.
Putnam, H. (1975a). "The Meaning of "Meaning'." in H. Putnam. *Mind, Language and Reality*. Cambridge: Cambridge University Press.
Putnam, H. (1975b). "Language and Reality." in H. Putnam, *Mind, Language and Reality*. Cambridge: Cambridge University Press.
Putnam, H. (1978). "Realism and Reason," in H. Putnam, *Meaning and The Moral Sciences*. London and New York: Routledge and Kegan Paul.
Putnam, H. (1980). "Models and Reality." *Journal of Symbolic Logic* **45**: 464-482.
Putnam, H. (1983). "Why There Isn't a Ready-Made World." in H. Putnam, *Realism and Reason*. Cambridge: Cambridge University Press.
Quine, W. V. O. (1969). "Natural Kinds." in W.V.0. Quine, *Ontological Relativity and Other Essays*. New York: Columbia University Press.
Railton, P. (1986). "Moral Realism." *Philosophical Review*. **95**: 163-207.
Sherman, P. and H. K. Reeve. (1997). "Forward and Backward: Alternative Approaches to Studying Human Social Evolution" in L. Betzig, ed. *Human Nature: A Critical Reader*. New York: Oxford University Press.
Sturgeon, N. (1984). "Moral Explanations." in D. Copp and D. Zimmerman (eds.) *Morality, Reason and Truth*. Totowa, N.J.: Rowman and Allanheld.
Sturr, C. (1998). *Ideology., Discursive Norms and Rationality*. Ithaca, New York: Ph.D. Dissertation, Cornell University.
Thornhill, R. and Thornhill, N.W. (1990). "An Evolutionary Analysis of Psychological Pain Following Rape: I. The Effects of Victim's and Marital Status." *Ethology and Sociobiology* **11**: 155-176.
Trivers, Robert L. (1971). "The Evolution of Reciprocal Altruism," Quarterly Review of Biology **46** (March 1971): 35-39, 45-47.
Williams, G. C. (1975). *Sex and Evolution*. Princeton: Princeton University Press.
Wilson, E.O. (1975). *Sociobiology: The New Synthesis*. Cambridge, Mass.: Harvard University Press.
Wilson, E.O. (1978). *On Human Nature*. Cambridge, Mass.: Harvard University Press.
Wood, A. W. (1972). "The Marxian Critique of Justice," *Philosophy and Public Affairs* **1**:

244-82.
Wood, A. W. (1979). "Marx on Right and Justice: A Reply to Husami," *Philosophy and Public Affairs* **8**: 267-95.
Wood, A. W. (1984). "A Marxian Approach 'to the Problem of Justice,'" *Philosophica* **33**: 9-32.