RICHARD N. BOYD

KINDS, COMPLEXITY AND MULTIPLE REALIZATION

Comments on Millikan's "Historical Kinds and the Special Sciences"

1. MILLIKAN'S CONCEPTION

1.1. *The Prevailing Picture*

According to Professor Millikan there is a prevailing picture of kinds in the special sciences (due to some temporal parts of Fodor) according to which:

1. Laws in the special sciences aren't exceptionless.
2. Kinds in those sciences are multiply realized.
3. #1 is at least partly explained by #2: the heterogeneity of the realizations explains why the laws of the special sciences are not reducible to the "proper law" of the physical sciences.

Psychology, with multiply realized functionally defined kinds, is the paradigm case to which this analysis is supposed to apply.

Millikan addresses a central question raised by this picture: Why are the laws of these sciences "proper laws" and not just accidental generalizations? Professor Millikan's answer is complex and subtle. The claims she makes on which I propose to focus are these:

1.2. *Kinds in the Special Sciences*

There are no special sciences *in Fodor's sense*, because:

1. Even though there are sciences whose laws are not exceptionless,
2. there are no sciences whose laws range over domains of kinds which are multiply realized in the sense intended by Fodor. There could not be any univocal *empirical* science which studies all of the different realizations of, e.g., functionally defined

Philosophical Studies **95:** 67–98, 1999.

psychological states. The various different realizations of such kinds are too diverse for there to be any *a posteriori* laws about them. They are not projectible, so

3. any "laws" applying to all instantiations of such functional kinds would not be *a posteriori* – they would follow just from the analytic functional definitions of the kinds in question.

1.3. *Historical Sciences and Historical Kinds*

The sciences, like psychology, whose laws are not exceptionless are *historical sciences* whose kinds – *historical kinds* – differ from eternal kinds in their definitional structures. Eternal natural kinds are natural kinds for which the ontological ground of induction lies in "the intrinsic nature of members of the kind" and not on "relative location in time and space."

By contrast, homeostatic property cluster kinds (HPC kinds) in the sense of Boyd 1991 are always historical kinds. The imperfection in the relevant property clustering occurs for such a kind because the property cluster is sustained over time by mechanisms of replication or information transfer between members of the kind so that members of a HPC kind:

1. are like one another, not because of a shared eternal essences, but because of historical relations between them;
2. exhibit the properties of the kind because the other members exhibit them;
3. support induction because such inductions are "grounded' because there is a certain kind of historical link between the members, for which reason the members are like one another; and
4. are modally spatially and historically delimited so that, for example, *Homo sapiens* could not occur on "Twin Earth."

Biological species are examples of this phenomenon as are historical artifact kinds (like the 1969 Plymouth Valiant): in both cases "copying" of information from historically situated tokens is the basis for the relevant homeostasis.

1.4. *A Non-Humean Approach to Laws and Generalizations*

Nevertheless, the "laws" about those historical kinds differ from "accidental generalizations" in that those kinds – just like eternal kinds – "... are answered in nature by a supporting ground of induction": causal structures which underwrite the inductive and explanatory practices of the relevant sciences. One should reject Humean conceptions which link the notion of non-accidental generalization with universal and exceptionless laws.

2. POINTS OF AGREEMENT

2.1. *Natural Kinds, Disciplines, Accommodation and Projectibility*

There is so much that I agree with in Professor Millikan's approach that I think my disagreements are best understood in the context of an account of the points on which we agree. According to Millikan the theory of natural kinds is essentially concerned with issues of objective projectibility. "Real Kinds" are "answered in nature by a supporting ground of induction" explicable, I gather, in terms of some non-Humean conception of how the causal powers of members of kinds cooperate somehow with the inductive practices of scientific disciplines. The naturalness of a natural kind depends on the disciplinary context within which it is employed – that's why functionally defined psychological states aren't natural kinds: they are so diverse in their realizations that there's no possible *a posteriori* discipline which studies them.

I've defended somewhat similar views about natural kinds and projectibility. I hold (Boyd, 1989, 1990a, 1990b, 1992, 1993) that what makes natural kinds *natural* – what makes reference to them contribute to projectibility judgments – is that reference to them allows us to achieve in *accommodation* between our classificatory and inductive and explanatory practices within a *disciplinary matrix* and the causal structures of relevant phenomena. This fit or accommodation between natural kind categories and induction-supporting causal powers of things is my version of what Millikan means when she says that natural kinds are "answered in nature by a supporting ground of induction." I have likewise argued (Boyd, 1982, 1983,

1989, 1993) that the reference relation between a natural kind term, t, and a kind, k, is to be understood in terms of ways in which the actual practices within disciplinary matrices bring it about that what is predicated of t tends to be (approximately) true of k. When this happens, I argue, one should think of the deployment of t in the relevant matrix as providing socially coordinated *epistemic access* to k.

By the *accommodation demands* of a disciplinary matrix I mean the sorts of accommodation between its classificatory and conceptual resources and relevant causal structures which would be required to provide (in Millikan's terms) a "supporting ground" for its inductive and explanatory practices. In particularly, I have it in mind that successful accommodation characteristically enables scientists to reliably discern which generalizations among those which fit available data are *causally sustained* in the relevant domains of application and hence will prove to hold true for future (or unexamined past and actual) cases. It is important for our present purposes to note that not all of the accommodation demands of a disciplinary matrix need be satisfiable: even in mature scientific disciplines there may be serious explanatory or inductive projects for which the causal structure of the relevant phenomena can provide no "supporting ground."

According to the position I am here developing there is a perfectly good sense of the term "definition" according to which a natural kind is defined by the role which the use of natural kind terms referring to it plays in satisfying the *accommodation demands* of a disciplinary matrix. Such definitions characteristically identify the functionally specified causal or explanatory role which disciplinary practices anticipate will be played by members of the kind in question. Call this sort of definition of a kind its *programmatic definition*. There is another perfectly legitimate sense of "definition" according to which the definition of a natural kind is provided by an account of the properties shared (perhaps imperfectly) by its members in virtue of which reference to the kind plays the role indicated by its programmatic definition. Call this sort of definition of a kind its *explanatory* definition.

To a good first approximation – what I mean by that is that I'm ignoring here the issues of partial denotation, non-referring

expressions, subtle questions about the individuation of disciplinary matrices, translation of natural kind terms between different languages employed within the same disciplinary matrix, etc. – I advocate the following conception of kind definitions and of reference:

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

1. (Epistemic access condition) There is a systematic, causally sustained, tendency – established by the causal relations between practices in M and causal structures in the world – for what is predicated of t_i within The practice of M to be approximately true of things which satisfy F_i, $i = 1, \ldots, n$.
2. (Accommodation condition) This fact, together with the causal powers of things satisfying these explanatory definitions, causally explains how the use of $t_1, \ldots, t_n$ in M contributes to accommodation of the inferential practices of M to relevant causal structures: to the tendency for participants in M to identify causally sustained generalizations and to obtain correct explanations.

Let $k_1, \ldots, k_n$, be the kinds referred to by $t_1, \ldots, t_n$. Then the proposition that they have $F_1, \ldots, F_n$ as their *explanatory definitions* explains why they satisfy the *programmatic definitions* determined by the accommodation demands of M.

2.2. *Laws and Sciences*

The conception of natural kinds just articulated permits me to explain the depth of my agreement with two of Professor Millikan's conclusions. In the first place, I too doubt that there could be a "single science" whose domain includes all of the realizations of some functionalist definition of psychological states. If natural kinds are just those which play a role in accommodation of disciplinary matrices, it is clear why there probably could not be a disciplinary matrix whose scope included research into laws governing all such realizations. The various realizations of, e.g., the functional definition of pain are probably so diverse that they have no regularity

sustaining causal properties in common, save only those mentioned in the functional definition.

Similarly, if natural kinds are those kinds reference to which is important in the formulation and recognition of *causally sustained* regularities of a sort appropriate to a disciplinary matrix, then it is easy to see why Millikan is right that laws – as opposed to accidental regularities – need not be exceptionless or universally applicable. Millikan says, "On a reasonable reading, a valid scientific law is just a true, well grounded, hence non-accidental, generalization – well grounded, that is, not just in logic but in ontology." Ontological grounding is, on the view I offer, a matter of causal sustenance, and there is no difficulty in explaining why there are disciplinary matrices (those involving historical disciplines, for example) which aim at the identification of causally sustained, but not exceptionless, regularities of less than universal scope.

2.3. *Species*

Similar considerations indicate why Professor Millikan is right to depart from current biological fashion and treat biological species as natural kinds. Natural kinds are solutions to problems of disciplinary accommodation: to problems about how to sort things so as to facilitate reliable induction and explanation. It is clear that all parties to philosophical disputes about whether biological species are kinds or individuals agree that the assignment of populations to species must meet such a requirement of accommodation. In particular, that's why species can't be defined as historically unrestricted categories characterized by necessary and sufficient membership conditions: no such way of sorting populations can satisfy the accommodation demands connected with the use of species names in biology, but other ways of doing so can.

Thus reference to biological species is just an instance of the accommodation of conceptual and classificatory resources to causal structures, which is what natural kinds *are for*. Once we get the metaphysics and epistemology of natural kinds right, we can see that, even if biological species are, in the relevant sense, historical individuals, their constituent populations constitute natural kinds nonetheless. So much for *analytic* metaphysics!

In fact, we can see that, from the point of view of the philosophical theory of projectibility and accommodation, the distinction between kinds and individuals is much less important that it might seem. Anyone who can define "grue" can define persistence conditions for gruish individuals which mix and match temporal stages of ordinary persisting individuals, or define otherwise *un*natural *things* as well as unnatural properties. Induction and explanation require that we quantify over natural as opposed to unnatural individuals every bit as much as that we deploy only projectible predicates. In each case what is at stake is whether or not the resulting inferential practices satisfy (in so far as they can be satisfied) the accommodation demands of the relevant disciplinary matrices. In a perfectly good sense (the one just described) the temporal stages and the spacial parts of a *natural* individual form a natural kind of sorts. Naturalness is not a matter of logical type or form, but of accommodation.

3. SOME POINTS OF DISAGREEMENT

3.1. *Analytic Functionalism, etc*

We routinely offer programmatic definitions of scientific kinds, of the sort mentioned in section 2.1, by describing in very abstract terms the causal or explanatory role we anticipate that such kinds (and reference to them) will play. The definitions of psychological states offered by various versions of "analytic functionalism" are examples of such programmatic definitions: psychological stats are characterized by very abstractly formulated causal or explanatory roles with respect to each other, sensory inputs and behaviors. I have just agreed with Professor Millikan that very often such there are no interesting "laws" or explanations about all of the nomologically possible instantiations of such programmatic definitions.

This point is independent of the conception – which Millikan seems to share – that such programmatic definitions are really *analytic* or otherwise a *priori*. I doubt that this is ever true of the sorts of causal/explanatory role definitions which philosophers take seriously. Here's why. Let E be the conjunction of a set of explanatory/causal role specifications which are (collectively) *prima facie* candidates for programmatic definitions of some natural kinds k_1,

$\ldots, k_n$, framed in terms of natural kind terms $t_1, \ldots, t_n$ referring to $k_1, \ldots, k_n$, respectively within the relevant disciplinary matrix, M. Let's call such a conjunction a candidate *conjunctive programmatic definition for M*. Let R be the Ramsey sentence of E, and, for each k_i, i = 1, …, n, let R_i be derived Ramsey sentence definition of membership in k_i. R and the various R_i's will involve quantifiers over natural kinds. [I leave it to the reader to persuade herself that the R_i's aren't even candidates for extensional correctness (or, perhaps, even for coherence) if the quantifiers range over arbitrary kinds – whatever those might be.]

Let us now define a *modest embarrassment* for E. E is *modestly embarrassed* just in case

(1) the true explanation of the satisfaction of the accommodation demands of M in respect of the deployment of $t_1, \ldots, t_n$ (in the sense articulated in section 2.1) assigns to $k_1, \ldots, k_n$ explanatory definitions $F_1, \ldots, F_n$ such that kinds so defined could not (given relevant laws of nature) have *exactly* the causal/explanatory roles assigned to $k_1, \ldots, k_n$ by E, but,
(2) E provides a revealing approximation to the central causal/explanatory roles of the kinds k_i.

A modestly embarrassed conjunctive programmatic definition, then, is an approximately true and revealing (albeit false) account of the causal and explanatory roles of a family of kinds which suit those kinds to the explanatory and inductive tasks of a disciplinary matrix. It's not a bad thing to propose, or even to believe, such a definition: doing so will get you close to the truth about how practices within the disciplinary matrix work.

It's also easy to see how even someone with an intimate first-hand knowledge of the aims, inductive and explanatory strategies, and the central methods of a disciplinary matrix could advocate – or become persuaded of – a modestly embarrassed conjunctive programmatic definition.

Suppose, for example, that you have an expert's appreciation of the aims and methods of psychological research on memory. You recognize that the distinction between "short term memory" and "long term memory" was introduced as part of the theoretical machinery for explaining certain real-life and experimental facts by positing one sort of information storage system (short term memory)

which retains information *prior to* its incorporation into certain more permanent information storage structures (long term memory). You incorporate exactly this conception into your programmatic definitions of long and short term memory and, thus, into your proposed conjunctive programmatic definition of the terminology of the relevant disciplinary matrix.

Now suppose that the deployment of the terms "short term memory" and "long term memory" does contribute to the accommodation of the matrix in question to relevant causal structures but that, as a matter of fact, when information is stored in the structures corresponding to the term "short term memory" that information is characteristically stored *simultaneously* in the structures referred to by the term "long term memory", but is not retrievable from the latter structures for some (relevantly long) time. Suppose, further, that memory systems are so integrated (as a result of evolved error-checking mechanisms, let us suppose) that when short term memory traces are erased before long term ones become accessible the latter get extinguished too.

If this is so, and the rest of your proposed programmatic definitions for psychological kinds are right, then the conjunctive programmatic definition you propose will be modestly embarrassed: it will provide an approximate and revealing characterization of the causal powers of, and causal relations between, psychological states upon which their suitability to the relevant causal/explanatory roles in psychological theorizing depend.

The associated Ramsey sentence will, of course, be false, since no natural kinds play just the causal roles assigned to the natural kinds of psychology by your proposed conjunctive programmatic definition. [The only candidates will be the natural kinds of the relevant matrix and they do not (not *exactly*) satisfy the matrix of the Ramsey sentence.] Likewise, the associated Ramsey sentence definitions for particular psychological kinds will each pick out the null set. The import of your mistake is magnified when you move to Ramsification.

Consider now the implications of what we have just seen for conceptions according to which some causal/functional role definitions of some natural kinds are analytic or otherwise *a priori*. Let $A(s_i, \ldots, s_m)$ be a proposed definition of this sort for natural kind

terms $s_i, \ldots, s_m$ employed within a disciplinary matrix N. If the terms in question refer, then there will be explanatory definitions, $G_1, \ldots, G_m$ which reflect the roles played by $s_i, \ldots, s_m$ in satisfying the accommodation demands of N. The Ramsey sentence definitions of $s_i, \ldots, s_m$ derived from $A(s_i, \ldots, s_m)$ will have non-null extensions only if the causal/explanatory roles specified by $A(s_i, \ldots, s_m)$ are *exactly* compatible with the causal powers determined by $G_1, \ldots, G_m$.

Thus, in order for $A(s_i, \ldots, s_m)$ and the associated Ramsey sentence definitions to be *a priori* it would have to be a truth knowable *a priori* that what $A(s_i, \ldots, s_m)$ asserts about causal/explanatory relations among the s_i's is *exactly* true of the kinds specified by the *a posteriori* explanatory definitions G_i which specify how accommodation is accomplished in a very complex family of the terms s_i. That sort of causal knowledge is, I suggest, never *a priori*.

This is not an esoteric point. Plainly a proposal which embodies a proposed explanatory definition of a natural kind is *a posteriori* since it embodies a detailed account of how accommodation is achieved in a particular matrix. A proposed programmatic definition embodies, instead, an account of the broad outlines of the basic structure of the *successful* causal explanations within such a matrix. It's harder for a sophisticated commentator to get it wrong about the latter sort of issue, but that hardly makes the issue an *a priori* one. Sufficiently substantial discoveries within a disciplinary matrix can dictate the rejection or modification of what had been the best supported programmatic definitions of the kinds it studies.

Thus, for example, the shift from what he calls "essentialist" to population thinking about taxa celebrated by Mayr (see, e.g., Mayr, 1988) can be best thought of as involving an *a posteriori* revision of prevailing programmatic functional definitions of biological taxa in the light of Darwin's discoveries. Once this sort of example is recognized, I believe, one can see that there is really something like a continuum between highly abstract functional programmatic definitions, like "Species are the basic units in evolution", which – although *a posteriori* – are not very informative and detailed explanatory definitions like "Water = H_2O."

What I think makes things confusing is that two philosophical tasks of identifying the "essential properties" of kinds have gotten conflated. On the one hand, we often feel (mistakenly, I have just argued) confident that we are able to identify *analytic* functional or causal/explanatory role definitions for various categories, and thereby to determine their essences. On the other hand, we are used to relying on a notion of essential properties according to which the essential properties of things in a natural kind are just those properties of things in the kind which underwrite the unity in the kind which makes it appropriate as a vehicle for explanation and induction. In the first sort of case, the definitions we have in mind will be very abstract programmatic definitions; in the second we'll have in mind what I have been calling explanatory definitions (or, at least, definitions near the explanatory end of the continuum).

What Professor Millikan's paper indicates is that the programmatic definitions we get in the first sort of inquiry do not line up nicely with the natural definitions we get in the latter. Her own treatment of the issue suggests that what is (partly) responsible for the mismatch is that the programmatic definitions are analytic.

In fact, I believe, this underestimates the extent of her insight. Even when definitions of kinds in terms of broad causal/explanatory roles are clearly *a posteriori* it may often – perhaps always for non-contrived cases – be the case that there could be no unitary science which studies all of the nomologically possible realizations of those definitions. I'll have more to say about this phenomenon later.

3.2. *Homeostasis, Historicity and Accommodation*

3.2.1. *A Clarification*

Professor Millikan makes several suggestions and criticisms regarding my articulation of the notion of a homeostatic property cluster kinds, and of my application of it in the case of biological species. Most of these I will deal with in the following (sub)sections, but I want to clarify some points before going on. Millikan says that, prior to my indicating that I thought that homeostatic property clusters are historically individuated, what I seemed to be proposing would have the consequence that HPC kinds would be a variety of eternal kinds because:

(1) the "lawful interdependence" of various surface/deep properties of an HPC kind would be the glue holding the kind together,
(2) the same HPC kinds might be found in other nomologically possible worlds so that, for example, *Homo sapiens* would occur on Twin Earth, and
(3) on that assumption (according to Millikan) it would remain unanswered why imperfect homeostasis is possible. She asks why homeostatic kinds are not "on the one hand, perfectly homeostatic", or on the other, "the results of large scale accident historical circumstances?"

In the spirit of these concerns Professor Millikan offers an interpretation of my account of HPC kinds and a correction to my application of it to the case of biological species. Remarking on my claim that HPC's are historically individuated she concludes that HPC kinds are, on my account, not eternal kinds after all but historical ones. She then asks what holds them together if not an eternal essence. She first remarks that the answer for species cannot be interbreeding to the extent suggested by my account, because there are known to be other sources of evolutionary stabilities.

She then notes, approvingly, that what the emphasis on interbreeding does is to limit each species to an historical location "*in this world*." She then asserts, on her own behalf, that "(b)iological kinds are defined by reference to historical relations among the members, not, in the first instance, by reference to properties .... So, they are historical kinds."

Millikan then generously offers a reformulation of my account of HPC kinds according to which (a) the sources of imperfect homeostasis lie in imperfections in replication or copying of (something like) information structures, (b) the need for historical individuation of HPC's arises from this sort of imperfect transmission, (c) HPC kinds are necessarily restricted to particular historical location determined by the relevant mechanisms of transmission, and (d) HPC kinds are defined by reference to historical relations among the members, not, in the first instance, by reference to shared properties. [At least I think the last point is what she intends for HPC kinds generally. It's what she defends for biological species and for 1969 Plymouth Valiants, and it's hard to see how, given her argumentative strategy, she would want to treat other HPC kinds differently.]

I very much appreciate Professor Millikan's sympathetic and careful reading of my efforts to articulate a conception of HPC kinds. I agree entirely with two of the points she makes. In the first place, I agree that, for all or almost all biological species, there are important mechanisms of homeostasis other than gene exchange, and I agree that this, means that, for some HPC kinds, some of the defining homeostatic mechanisms are external to the entities that fall within the kinds. Indeed I *intended* to convey both these points in my earlier efforts to describe HPC kinds. In particular, I intended my references to reproductive isolation and gene exchange simply to indicate that an appreciation of one sort of homeostatic mechanism was part of one standard approach to the paradigm cases of natural kinds. I didn't intend to subscribe to Mayr's biological species definition, but simply to refer to it, and to indicate that, for those cases in which it provides some insight into what defines a species, it validates the HPC conception. Every commentator I have read (or spoken with) about what I wrote has taken me to endorse roughly Mayr's conception, so obviously the problem lies in the unclarity of my exposition.

I dissent from the other conceptions which Millikan advances about HPC kinds. Although I think that HPCs are *prima facie* subject to historical standards of individuation, I do not believe that in all cases there will be spatio-temporal variability in their constituent properties and mechanism. Thus I do not intend to always explain the possibility of imperfect homeostasis by appealing to spatio-temporal variability in underlying HPCs. Indeed, I don't think that this will typically be the most important explanation even in cases in which there is such variability in the relevant HPC. I intend the general characterization of HPC kinds to be neutral with respect to all of these issues.

I likewise intended the account of an HPC kind to be neutral with respect to whether mechanisms of copying or information transmission are important to homeostasis or to its imperfections. As I'll argue, there are clear cases of HPC kinds where no such mechanisms operate.

One issue on which I did not intend to be neutral is this: Even when the relevant homeostatic mechanisms crucially involve copying or information transfer – as in the case of biological species –

I do not, for better or worse, hold that HPC kinds are defined by reference to historical relations among the members, rather than by reference to their shared properties. It is to this and related issues that I now turn.

3.2.2. *Species*

Professor Millikan maintains that biological species are limited to particular historical locations and that, in the first instance, they are defined by historical relations between their members rather than by the property cluster their members have in common. It is widely agreed among biologists and philosophers of biology that species are historically delimited. If we suppose that in some contemporary pond there happened to be a population of organisms physically identical in every respect to a population of some early Jurassic fish, they would not be members of the same species. The reason, it is agreed, is that external factors are so crucial to the evolutionary fate of a species that the contemporary organisms and their Jurassic analogues would not share the same "evolutionary fate" to the extent appropriate for conspecificity.

Often it is concluded from such considerations that the populations within a species must be parts of a single lineage. This seems to be Millikan's view when she holds that species are in the first instance defined by historical relations between their members – and, by analogy, when she is at pains to insist that all 1969 Plymouth Valiants are descended, as it were, from a single plan.

I agree that biological species are necessarily limited to a particular historical situation, but I deny both of the further conclusions. In the first place, there are actual cases of recognized and evolutionarily stable plant species which arose from hybridization between other distinct species. Certainly such species *can* evolve – and almost certainly they *have* evolved – more than once giving rise to distinct lineages within the same species. It remains true in such cases that the commonalities in evolutionary tendencies we expect of conspecific populations require that independently evolved conspecific lineages of this sort arise in within the same historical framework. Similarly, the stability of such a hybrid species over time will typically (perhaps always) be partly explained by the sorts of historical relations Millikan has in mind *among members of the "parent"*

species. But all this is compatible with the independent evolution of two or more lineages within the hybrid species.

A similar consideration of the necessary historical situatedness of biological species also indicates why we should be reluctant to agree with Millikan that their definition is primarily a matter of historical relations among their members rather than of shared properties. Species should be individuated historically, the argument goes, so that reference to species affords us the right sort of resource for evolutionary explanations. There must be a tendency for populations within a species to be subject to approximately the same evolutionary forces, and this would not be so for populations under sufficiently different historical circumstances. Quite so; *nor would there be such a tendency if members of the same species (of the same sex and degree of maturity) did not share a great many phenotypic characters*. Species are defined, according to the HPC conception, by those shared properties *and* by the mechanisms (including both "external" mechanisms and genetic transmission) which sustain their homeostasis.

In fact, we take these shared characteristics for granted in ordinary evolutionary explanations. Any evolutionary explanation – whether adaptationist or neutralist – of any phenotypic trait characteristic of a species takes for granted an almost unsurveyabily large number of phenotypic similarities and the homeostatic mechanisms which tend to sustain them over time. A biological species is a locus of a high level of evolutionary stability, not just a lineage linked by descent.

A quite different set of considerations leads to the same conclusion about the definitional importance of shared homeostatically linked properties. One of the defining features of natural kinds generally (one of the features which makes natural kinds form a *natural kind* in philosophy) is that reference to natural kinds facilitates induction and explanation with respect to a wide variety of issues – often beyond the domain of a single scientific discipline as these are ordinarily understood. The latter is true, for example, of the natural kinds of chemistry.

It is important to remember that something similar is true of biological species. Reference to them plays a central role induction and explanation, not just in evolutionary biology, but in the

rest of biology as well. Only the (imperfectly) shared properties characteristic of biological species can explain how the relevant accommodation of classificatory practices to causal structures is achieved. It is also true of biological species, as Millikan indicates it is of many other natural kinds, that their members are often *pretty* uniform, so that studying a relatively small sample from a species gets information about (almost) all of its members. Of course the variability within species makes such inferences less secure than such inferences in chemistry, but the same is true of inferences in, e.g., geology and meteorology. Here too, it is the shared features rather than historical relations which are centrally important.

I think that it is easy to underestimate the importance of shared properties here if one places the wrong emphasis on the important fact that evolutionary biology is an historical science. Even the historical explanations it provides depend importantly on tacit (or explicit) reference to those shared properties. It is an important fact that the stability of the property cluster associated with a particular species over time (and space) is entirely a product of historical relations between populations together with external historical contingencies. In this sense species are essentially historical entities. It does not follow, however, that these historical phenomena *and not the property homeostasis they sustain* form the essence of the explanatory definition of a species.

3.2.3. *Non-Historical Homeostatic Property Cluster Kinds*

Professor Millikan suggests that the source of imperfect homeostasis in a HPC kind must be something like faulty replication from one historical stage to another, and that HPC kinds must therefore be necessarily restricted to some appropriate historical location. While I agree (with the qualifications just discussed) that this is so for biological species, I think that the HPC analysis applies to lots of natural kinds about which neither of these claims is true. Meteorological classification of storm patterns, for example, or the classification of minerals into what mineralogists call mineral *species* are examples. In each case, there are certain meteorological or geological mechanisms which tend to the production of phenomena (storms and tokens of mineral types) exhibiting (imperfectly) a characteristic cluster of scientifically interesting proper-

ties, and the kinds (of storms and of minerals) whose recognition affords us accommodation to the relevant meteorological or geological causal factors certainly seem to be defined by the relevant property clusters, together with the mechanisms characteristically responsible for their co-occurrence. Certainly explanations and non-accidental inductive generalizations about three phenomena depend for their causal appropriateness on just the sorts of causal factors the HPC conception identifies.

Of course, in both these cases, the imperfections in homeostasis are to be explained by historical contingencies associated with the production of particular storms or mineral tokens. But in neither case is what is at stake anything like historical causal relations of any sort *between* instances. Nor are such kinds properly thought of as necessarily restricted to particular historical locations. Meteorological and geological kinds almost certainly depend for their possible realizations on quite particular conditions which may be peculiar to Earth. Nevertheless, the fact that the relevant homeostatic mechanisms do not involve anything like historical transmission means that if relevantly similar conditions did obtain on "Twin Earth", good methodology would require us to deploy or current meteorological and geological categories to them, just as we now sometimes now properly apply the same mineralogical term to specimens from very different locations and geological times.

Once you're on a roll with these sorts of examples I suspect that you can find lots more. Types of galaxies and of star clusters are plausible candidates; so are types of geological formations. In the social sciences, it is plausible that certain economic categories – capitalism, for example – are HPC kinds which are not historical in the sense articulated by Millikan. Perhaps the commonalities imperfectly shared by capitalist economies are all to be explained by patterns of common descent from early trading activities in medieval European towns, or something of the sort, but this is not implied merely by the HPC character of capitalism as a natural kind in economics.

Even in chemistry there may be examples of non-historical HPC kinds. The classification of elements into metals, semi-metals and non-metals has many of the features of HPC classifications. There is a cluster of metallic properties (thermal and electrical conduc-

tivity, ductility, malleability, and an inverse relationships between conductivity and temperature, for examples) which approximately increase as one moves down and to the left in the periodic table, and a contrasting set of non-metallic properties. Semi-metals (imperfectly) exhibit intermediate values of these traits. The imperfection of clustering in these categories is reflected in some variability in the assignment of elements to these categories, as you can see for yourself by looking up "semi-metals" in several different chemistry books. Now perhaps these categories do not represent full-blown natural kinds, but reference to them certainly contributes to accommodation, and they exhibit HPC-like characteristics without being historical kinds at all.

I conclude that natural kinds primarily defined by historical relations between their members (if there are any) form a paper sub-set of the HPC kinds and that this sub-set does not include biological species, even though they are paradigmatically *historical* in important senses of the term.

3.3. *Modal Properties of Natural Kinds*

3.3.1. *Eternal Kinds and Intrinsic Definitions*

I suspect that one reason Professor Millikan associates HPC kinds with historical kinds – both in the sense of kinds necessarily limited to an historical region and in the sense of kinds defined in the first instance by historical relations between members – is that she contrasts *eternal* natural kinds, defined by "central *intrinsic* properties, say from an inner structure common to all members of the kind", with *historical* kinds in the senses just mentioned, apparently taking this classificatory scheme to be exhaustive. Since biological species and lots of other HPC kinds are defined partly by *extrinsic* properties, they will all then appear to be historical in both senses.

If by an eternal natural kind one means a natural kind which is historical in neither of the senses just mentioned, then many meteorological kinds – and perhaps some astronomical kinds as well – may turn out to be partly *extrinsically* defined but still *eternal* kinds. The kind of stability which defines a natural kind of storm system, for example, may depend, unexpectedly perhaps, on the nature of weather systems distant from the storm itself. Perhaps distinctive sorts of interstellar dust clouds depend for their distinctive properties

partly on their relation to the distribution of nearby stars and their gravitational fields. In any event, there is no good reason to suppose that the essences – that is, the explanatory definitions – of eternal natural kinds must always consist only of intrinsic properties.

If by an *eternal* kind one means a kind eternal in the sense just discussed which also plays a role in *exceptionless* laws the situation is more complex. Exceptionless laws may be hard to come by; perhaps the only ones are those applying to the most basic elementary particles and their physical properties (assuming that there are most basic elementary particles). *Perhaps* it is true of them that the natural kinds into which they fall all have purely intrinsic explanatory definitions.

Perhaps. But even with respect to this notion of an eternal kind, the burden of proof should lie, I suggest, with the defender of the view that their essences are always intrinsic. It (epistemically speaking) might have turned out that atoms enter into exceptionless laws (indeed, they probably do). It might have turned out (but apparently it didn't) that the integrity of an individual atoms, upon which the truth of the exceptionless laws applying to it always depends, itself depends on the causal relations between that atom and others nearby. If this had been so, there would have been an extrinsic relational property which was part of the explanatory definition of every atomic natural kind.

It is a deep fact about nature that atomic kinds do not have partly extrinsic essences. It would be a deeper fact still that no eternal natural kind (in the present sense) has such an essence; no extant philosophical (or scientific) argument seriously addresses the possibility that we might someday discover a counterexample.

This is easier to see if we consider one way in which a certain sort of counterexample might arise. One standard explanation for nonlocal phenomena in quantum mechanics has it that some quantum mechanical systems (like those involving pairs of jointly produced particles with net spin zero) retain a holistic integrity as quantum systems even when their constituents, as we ordinarily understand them, are widely separated, and that certain causal interactions which intuitively *seem* to involve only one constituent (a z-spin measurement for example) *actually* involve an instantaneous change

in the entire system (determining, for example, the z-spin of another constituent).

Suppose that further investigations led to the confirmation of something like this picture, with the additional feature that no locally specified physical system fails to enter into theoretically important membership in such a widely distributed system. Were this so, then fundamental physics (where – we are supposing – exceptionless laws, if there are any, are to be found) might have the consequence that, in the final analysis, the notion of a property strictly intrinsic to a physical system is unavailable, and would need to be replaced by some more complex and currently unanticipated notion. Were this so, every locally characterized physical natural kind might be a counterexample to the proposal we are considering.

I am not suggesting that the possibility I just discussed will come to pass. What I do suggest is that we are not in a position to say that nothing relevantly like it is true. For all we know there may be lots of eternal natural kinds which figure in exceptionless laws but have partly extrinsic essences.

3.3.2. *Natural Kinds in Weird Worlds*

In explaining her conception of eternal kinds as "held together by universal and eternal laws of nature that determine the various properties of the kind from central intrinsic properties", Millikan says about eternal kinds that they thus "have all of their various properties of necessity." It's not exactly clear what she has in mind, but she seems to be endorsing the widespread view that natural laws (at any rate natural laws about eternal kinds) are necessary (but of course *a posteriori*) truths: that they hold in all possible worlds. I think that this is mistaken, but that's not the point I want to make here. Instead, I want to argue – in a way which seems to me congenial to Millikan's overall project – that it's a mistake when talking about natural kinds to inquire about their members, or the laws they obey, "in all possible worlds."

I argued in section 2.1 that the philosophical theory of natural kinds has, as its only subject matter, the ways in which the accommodation demands of various disciplinary matrices are, or could be, satisfied. I indicated how the explanatory definitions of the kinds referred to within a disciplinary matrix are to be understood in

terms of the satisfaction of epistemic and accommodation conditions formulated in terms of the accommodation demands of the matrix in question. Natural kinds, on that view, get their essences from the ways in which accommodation is actually accomplished.

There are deep issues here that go beyond the scope of these remarks (for a fuller account see Boyd forthcoming), but I believe that the centrality of considerations of accommodation to the theory of natural kinds dictates a particular conception of how the extension of a natural kind is determined in a non-actual possible world. Since this conception also rationalizes most of our philosophical practice regarding natural kinds in diverse possible worlds, there is independent reason to believe that it is true. Here's the conception I favor:

Let $t_1, \ldots, t_n$ be the natural kind terms deployed in an actual world disciplinary matrix M and let $F_1, \ldots, F_n$ be the explanatory definitions of the corresponding natural kinds, $k_1, \ldots, k_n$, satisfying the epistemic access and accommodation conditions of section 2.1. Let W be some non-actual world. Then, if $k_1, \ldots, k_n$ exist in W, their explanatory definitions ${}^{w}F_1, \ldots, {}^{w}F_n$ *in W* will be just those families of properties which satisfy the corresponding epistemic access and accommodation demands for the disciplinary matrix M *as it would have to be implemented regarding W*.

So, an object x will, in W, be in k_i ($i = 1, \ldots, n$) just in case so classifying it is central to its being the case that causal relations among things so classified *in W* will result in the satisfaction of the accommodation demands of M *for W*. ${}^{w}F_i$ will differ from F_i in just whatever ways are required to preserve the sort of accommodation achieved in the actual world through the use of t_i in M, given the ways in which W differs in its causal structure from the actual world.

Here's the *basic* idea: membership in a particular natural kind is a complex causal capacity defined with respect to the accommodation demands of an actual-world disciplinary matrix; to be a member of a kind in some non-actual world is to have the right causal capacities with respect to the accommodation demands of the *very same* disciplinary matrix, understood as an explanatory and inductive project implemented *in that other world*. This conception seems to answer to our standard philosophical practice: asked what the definition of, say, a biological natural kind would be in some

possible world, we imagine ourselves doing biology in that world and ask what definition would be appropriate. [Note: In choosing an example involving biological kinds in other possible worlds I am not going back on the idea that such kinds are limited to particular historical times. Non-actual possible worlds include possible variations on the history of the actual world. The sentence, "If it had rained for ten minutes more in Oberlin last night, *Felis catus* would still have survived as a species", is true even though (on the standard analysis) it says that a biological species exists in a non-actual (but very nearby) possible world.]

In the account I offer, instead of maintaining that the explanatory definitions of kinds in a non-actual possible world are fixed by the accommodation demands of the relevant disciplinary matrix *in* that world, I speak instead of the matrix in question being implemented *for* or *regarding* that world. We often evaluate claims about natural kinds in possible worlds in which there could not be any disciplinary practices at all. For example we know that the natural kinds of chemistry would have the same definitions in any possible world just like the actual one except that it contains no cognizing systems – and thus no disciplinary practices. In such cases what we do in practice is to envision implementing the inductive and explanatory aims of the relevant disciplinary matrix from a position which somehow affords us a birds-eye view of the non-actual world without actually being *in* it. Exactly how this abstraction is to be understood (see Boyd forthcoming) is less important than the fact that even in such cases the explanatory definitions of natural kinds are determined by the accommodation demands of actual world disciplinary matrices reconfigured to fit possible world in question.

It remains to see why we should be suspicious of efforts to determine what properties natural kinds have *in all possible worlds*. The first thing to notice is that in the case of a HPC kind, the explanatory definition is provided by a (perhaps historically individuated) *process* of homeostatic property cluster*ing*. In a non-actual possible world an object is in such a kind just in case it relevantly participates in the *very same* process in that world. But, the individuation of historical processes – like individuation of other historical entities, like persons, wars, nations, etc. – breaks down as one get very far from the actual world. For more distant worlds

it will often be indeterminate whether they contain the very same homeostatic process as that which defines a HPC kind in the actual world, so it will be indeterminate whether or not the kind exists in the world in question (for a more extensive discussion see Boyd, 1988).

We may put the point about HPC kinds this way: Think of natural kinds as being established by a sort of bicameral linguistic legislation in which we and the world jointly legislate. Our legislative role consists of implementing disciplinary matrices with their associated accommodation demands. The legislative role of the world consists in determining how and to what extent those demands can be met. Together we thereby establish the explanatory definitions of natural kinds. What we have just seen is that, in the case of HPC kinds, our bicameral legislative authority is limited: the sort of legislation we (jointly) engage in establishes definitions only for rather nearby possible worlds.

What I now want to indicate is that similar considerations show that our legislative authority is limited even with respect to non-HPC natural kinds. We have already seen that a crude but illustrative way of articulating the conception defended here is this: one must classify biological organisms in a possible world, W, just as *biology done in W* would require; and similarly for physical kinds and physics, for chemical kinds and chemistry, etc. The deficiencies of such a formulation lie in the facts that (a) there is no good reason to believe that the disciplinary matrices within which natural kind terms function are restricted to – or map neatly onto – disciplinary category terms like "biology" and (b) one would really need to talk of "biology (or physics or chemistry) done *for* W", for the reasons just rehearsed.

Even so, the formulation just offered allows us to see an important point about the definitions of natural kinds in non-actual worlds. On the simplified formulation the extension of a biological kind, k, is determined in a non-actual world, W, only if there is a sufficiently determinate answer to the question, "What is the manifestation *in W* of biology?" When we move away from the simplified formulation we see that, in so far as the analysis specifies membership conditions for a natural kind, k, in a non-actual possible world, W, it does so only if there is a sufficiently determinate answer to the

question "What is the manifestation *for W* of the disciplinary matrix within which reference to k is central in the actual world?" A determinate extension for a natural kind in a non-actual possible world, W, depends on their being, for W, a suitably determinate manifestation of the relevant actual world disciplinary matrix *appropriate for induction and explanation about W*.

Now actual world disciplinary matrices are families of social and instrumental practices which themselves have homeostatic property cluster features. They involve the simultaneous satisfaction of a very large number and wide range of accommodation demands. It is this pattern of simultaneous satisfaction of accommodation demands through the deployment of a suitably accommodated vocabulary which the theory of natural kinds aims to explain. It is characteristic of natural kind terms that, when their use is suitably accommodated to causal structures, the satisfaction of some of the accommodation demands of a matrix is conductive to the satisfaction of the others. This homeostatic unity of instances of accommodation demand satisfaction is what defines a discipline or disciplinary matrix and the kinds distinctive to it.

Consequently, when we individuate disciplinary matrices within or for various possible worlds we are individuating HPC phenomena, and for possible worlds sufficiently distant from the actual world it becomes indeterminate what the manifestation is for that world of any given actual world disciplinary matrix. In such cases the explanatory definition of the relevant natural kinds will likewise be indeterminate. In so far as natural kinds and kind terms go, we (and the actual world) are simply not capable of legislating "for all possible worlds."

It seems to me that we often operate philosophically as though the semantics of the languages we speak underwrites truth values for a very wide variety of counterfactuals about natural kinds in very weird possible worlds. Especially when we discuss issues about multiple realizability, we often consult our intuitions about quite bizarre cases indeed. If what I have just said is true, in most cases we are consulting our intuitions about questions which lack answers. It seems to me one of the merits of Professor Millikan's discussion of multiple realizable natural kinds that the arguments she deploys do not depend in this way on weird examples: she is concerned largely

with the naturalness of kinds in nomologically possible systems and with the demands of projectibility in the actual world.

I want to plead that we should in general follow her example and not focus on weird cases while we are trying to get things straight about multiple realizability as a matter of methodological concern in science; we do not, after all, practice science in any possible world but this one. Let's free discussions of realization, reduction, emergence, supervenience and the like from reliance on very far out examples and counterexamples.

4. MULTIPLE REALIZATION, DISCIPLINARY PLURALISM, AND NATURAL KINDS

4.1. *Grades of Multiple Realizability*

There are multiple realizations and multiple realizations. Your liver is multiple realized: its exact physical realization varies from moment to moment, and human livers generally are to a greater extent multiply realized. Then there is the realization in metaphysically possible ectoplasm of a Turing machine functional definition of "pain." There is no science which studies just your liver, nor *could* there be one which studies both human and ectoplasmic pain. But there are sciences (human anatomy and physiology) which study human livers, and within which the term "liver" functions as a successful natural kind term. In the case of the category of human livers, multiple realizability is not a barrier to natural-kindhood.

Let's call the first sort of multiple realizability *unproblematical*. Professor Millikan is aware of the distinction between problematical and unproblematical cases of multiple realizability. To a good first approximation, she seems to hold that, for psychological kinds at least, unproblematical multiple realizability is limited to the sort which occurs within a single species. In particular, she rejects an argument by Papineau and Macdonald according to which multiply realized functional states can fall within the purview of a single science when they all result from the same sort of selection for mediating particular relations between environmental inputs and outputs of the systems in which they figure.

This proposal would perhaps work, she suggests, if the laws governing the alleged interspecific kinds in question could be

thought of as reflections of the causal powers of evolutionary niches, conceived as phenomena independent of the organisms which occupy them. She then argues that niches cannot be so conceived because they are characteristically phenomena which organisms create as they evolve.

This is an extraordinarily subtle and interesting rebuttal. I'm not at all sure whether or not there is a cogent reply available to Papineau and Macdonald. What I do believe is that there are many cases in which a functionally defined kind *understood as realized in a wider range of species* (but not ectoplasmically!) functions as a genuine natural kind in a unitary scientific discipline. I think the explanation I'll offer for such cases will prove generally congenial to Millikan's own projects and may represent what Papineau and Macdonald should have said.

4.2. *Disciplinary Pluralism and Interspecific Kinds*

What I have in mind as cases of inter-species natural kinds are (once again) livers, and also psychological states. There is a unitary scientific discipline (human anatomy) which studies, among other things, human livers and with respect to whose accommodation demands human livers form a natural kind. But there is a also a discipline (mammalian anatomy) which studies, among other things, mammalian livers, and with respect to which they form a natural kind. Then there is a vertebrate anatomy . . . You get the point. [But note that there isn't a unitary discipline of anatomy for all possible creatures, or even for all possible liver-bearers.]

Now it is true that, with respect to certain parameters (those relevant to the anatomy of any particular vertebrate species), there are many fewer non-accidental generalizations about vertebrate livers than there are about, say, human livers. But inter-specific comparative and evolutionary questions arise in vertebrate anatomy which do not arise with respect to the anatomy of any single species. With respect to the inductive and explanatory aim of investigating *these* questions, the category of vertebrate livers is a paradigmatic natural kind.

Of course, all I am appealing to here is a kind of disciplinary or sub-disciplinary pluralism about natural kinds: some kinds are natural kinds because reference to them is suited to meeting the

accommodation demands of some disciplines; others are natural kinds because reference to them is suited to meeting the accommodation demands of other disciplines – and this can be so even when in some sense or other the disciplines are all subdisciplines of some broader discipline like biology. For some important sub-disciplines of biology inter-specifically multiply realizable functionally characterized kinds *are* natural kinds.

Of course what this suggests – to me at any rate – is that we can see why there could be human psychology, primate psychology, perhaps mammalian psychology, vertebrate psychology or whatever: different empirical sciences, not each studying the psychology of a different species, but instead each corresponding to a different level of abstraction in multiple realizability of psychological states, and addressing *different* sorts of questions. It's an empirical question which of these could be a unitary science, but given the evolutionary, genetic and developmental commonalities across the categories of animals in question, it is by no means obvious that they could not *all* be unitary disciplines. [Note again that I have not argued that there is any reason to believe that there could be a unitary discipline which studies *the psychology* of vertebrates, Martians, connectionist machines, and . . .]

I suspect that Professor Millikan would find these example *prima facie* congenial but would rightly wonder whether it would be possible to offer an account of which the categories of species are over which inter-specific unitary sciences of functional kinds are possible, without resorting to the arguments of Papineau and Macdonald which she rejects. The answer is "yes."

4.3. *Replacement Stability and Unproblematical Multiple Realization*

4.3.1. *Individual Objects*

Consider the persistence of an individual object or the fact that it persists in having some property – let's say shape – over some interval of time. Now the micro-realizations of the object – and of its having the shape it persists in having – are constantly changing in little ways. Objects and their token instantiations of shape properties are process-like phenomena, and different micro-constituents are involved from moment to moment. A certain sort of multiple

realizability of objects and their token properties is part of what it is for the objects to have their characteristic integrity over time. There are causal mechanisms – mechanisms central to the persistence of the object – whose operation guarantees (over a suitable range of conditions) that one of these micro-realizations is always replaced by another.

Such objects and their properties exhibit a distinct kind of multiple realizability guaranteed by mechanisms ensuring a certain sort of stability in micro-realizations: *replacement stability* let's call it. For a *certain range* of the logically possible realizations of any actual ping pong ball (*not* including, for example, the possible micro-realizations resulting when logically possible Martians substitute microscopic spaceships piloted by miniature Martians for some of its molecules) there are causally sustained (that is: non-accidental) tendencies in nature for one of these micro-realizations to be replaced by another. It is the operation of the underlying replacement processes which constitute the persistence of the ping pong ball over time. Moreover – and this is important – it is because these processes operate in ping pong balls (and, thus, because they exhibit the associated levels of multiple realizability) that they have the stable macroscopic properties (like shape and coefficient of restitution) which makes them fall under the non-accidental generalizations of Newtonian mechanics. The *naturalness* of individual ping pong balls from the point of view of mechanics thus depends on the causal processes which given each of them a certain level of multiple realizability.

4.3.2. *Replacement Stability and Kinds*

I think that a similar sort of replacement stability can be discerned for kinds, categories and properties themselves. Your liver and its physiological properties exhibit multiple realizability ensured by replacement stability but so does the category "human liver." Human livers are multiply realizable, but there are processes which tend to ensure a certain non-accidental stability in their realizations *above and beyond* the replacement stability manifested in individual livers and their properties. There are genetic and developmental mechanisms which tend to ensure that all humans will exhibit important commonalities in their livers' anatomical and physiological proper-

ties – even though the range of viable human livers is much greater than the range of viable liver manifestations in any single human. These commonalities are responsible for its being the case that human livers form a natural kind in human anatomy.

Thus, just as in the case of individual livers, the multiple realizability ensured by the operation of replacement stability producing processes is partly constitutive of the naturalness, for anatomy, of the kind *human liver*.

Likewise, there is multiple realizability – underwritten by processes guaranteeing replacement stability – with respect to the category "mammalian liver", and perhaps for some still wider taxonomic categories. For none of these categories, however, does the range of possible liver micro-realizations involved extend to all logically possible micro-realizations of the functionally specified category *liver*.

What I propose is that when multiple realizability is thus a matter of replacement stability there is the possibility of a univocal science concerned with all of the relevant micro-realizations, provided that the commonalities produced by the relevant replacement stabilizing processes are sufficiently robust and relevant to accommodation. *Human anatomy* is a unitary discipline, so is *mammalian anatomy*, so might *vertebrate anatomy* be. The viability of such sciences would depend on what sorts of causally stable uniformities, partly guaranteed by the operation of processes of replacement stabilization, obtain across the relevant species. As one moves up in levels of abstraction from your liver, to human livers, to mammalian livers, etc., livers become less uniform, but new laws may "emerge" underwritten, for example, by facts about the evolution of vertebrate anatomical systems.

This means, by analogy, that we can see why there could be human psychology, perhaps mammalian psychology, vertebrate psychological or whatever: different empirical sciences not studying he psychology of different species but corresponding to different levels of abstraction in multiple realizability of psychological states and to different sorts of laws.

Why not the psychology of a category of creatures so wide that it includes ectoplasmically realized Martian in a logically possible world in which Mars is made of non-material green cheese or, at

least, in whatever extra-terrestrial cognizers there may be in the actual world? Answer: So far as we know, there are not processes of replacement stabilization linking their properties to ours.

I indicated that I thought that Professor Millikan might find this proposal congenial. She would, in any event, want to ask how it differs from the proposal of Papineau and Macdonald which she rejects. Here's the answer:

Papineau and Macdonald's proposal suggests that a functionally characterized and multiply realized kind can form a natural kind if certain quite specific adaptive processes serve to establish relevant similarities between its members. Such processes would be replacement stabilizing processes in the sense just discussed. Professor Millikan argues that the processes in question could only produce similarities capable of supporting "occasional illuminating comparisons across species" rather than genuine laws. What appears to be the case is that Professor Millikan would accept as natural kinds inter-specific multiply realized functional kinds for which replacement stabilizing processes in fact produced sufficiently causally important commonalities among their members.

What I have been suggesting is that, for livers, and probably for various psychological states as well, there are such stabilizing processes. What I have not done is to require or suggest that they must all be the particular sorts of adaptive processes considered by Papineau and Macdonald. I would suppose that, in addition to selective forces, there would also be stabilizing factors arising from inter-specific similarities in genetic structures and from developmental constraints leading to important neurophysiological and neurochemical similarities. Should this prove to be so we would, I think, have – by Millikan's standards as well as by mine – instances of unproblematical inter-specific realization of the sort at issue.

4.3.3. *The Species Category and a Note on Higher Taxa*

Individual biological species are, of course, multiply realized and the stabilities they exhibit in biologically important properties are examples of replacement stabilities. The category *biological species* is itself (spectacularly) multiply realized and (probably) has some informative functional characterization. It is worth asking whether

it is a natural kind in evolutionary biology. This is a somewhat odd question: it's like asking whether or not the kind *mineral* is a natural kind in mineralogy. In neither case would a negative answer provide a methodological reason to abandon the category. Still, no one would expect the notion of a mineral to play a central role in very many laws, but, I suspect, the issue seems less clear with respect to species – especially if one adopts (as one should) the non-Humean approach advocated by Millikan, so that one does not require of laws that they be ahistorical and exceptionless.

It seems to me that one way to approach this question is suggested by Professor Millikan's insistence on the variety of homeostatic (= replacement stabilizing) mechanisms which can operate to preserve the integrity of a species over time. It is reasonable to suppose that, if the category species does not form a natural kind in evolutionary biology, the reason it does not lies in the different mixes of these mechanisms which sustain homeostasis in different sorts of species.

If this proves to be so, then it would be important to inquire whether there are natural sub-categories within the category *species* defined in terms of different configurations of homeostatic mechanisms. Perhaps members of these sub-categories would share a common "ontological ground" for induction or explanation in evolutionary biology, and thus be genuine natural kinds.

I can't resist one more Millikanesque speculation about natural kinds in evolutionary biology. According to many cladists higher taxa must be strictly monophyletic. Their reasons for insisting on this principle vary. Sometimes, for example, those reasons rest on what are essentially positivist critiques of theory-dependent schemes of classification in general.

The most impressive cladist arguments turn on critiques of "adaptationist" conceptions in evolutionary theory. One influential argument for the alternative "evolutionary systematic" approaches to higher taxa takes polyphyletic taxa to be justifiable when they contain closely related organisms displaying the same adaptive "evolutionary innovation." In such cases, this argument proposes, they will tend to exhibit similar responses to selection pressures, and thus exhibit similar evolutionary tendencies. This similarity justifies us in classifying them together. For some cladists who assign a less

important role to natural selection in evolution, this argument has proven unconvincing (see Guyot, 1987).

What I think would be interesting would be a reworking of this sort of argument for polyphyletic taxa in which a wider range of similarity sustaining mechanisms are considered, including, for example, non-adaptive developmental constraints on variability within related lineages. On this alternative conception, higher taxa (including some polyphyletic ones) would be natural kinds defined in terms of phyletic inertia, broadly understood, rather than in purely adaptationist terms. I haven't the foggiest idea whether or not such a argument could ultimately be sustained, but it seems to me that it would be a natural extension of Professor Millikan's insistence on the variety of different homeostasis sustaining mechanisms in the case of individual species.

REFERENCES

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

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. (forthcoming): 'Kinds as the "Workmanship of Men": Realism, Constructivism, and Natural Kinds', *Proceedings of the Third International Congress, Gesellschaft für Analytische Philosophie*. de Gruyter.

Guyot, K. (1987): *What, If Anything, is a Higher Taxon?* Ph.D thesis Cornell University.

Mayr, E. (1988): *Towards a New Philosophy of Biology*, Cambridge: Harvard University Press.

Cornell University