

This excerpt is provided, in screen-viewable form, for personal use only by members of MIT CogNet.

Unauthorized use or dissemination of this information is expressly forbidden.

If you have any questions about this material, please contact [cognetadmin@cognet.mit.edu](mailto:cognetadmin@cognet.mit.edu)

## Squaring the Circle: Natural Kinds with Historical Essences

Paul E. Griffiths

### FROM ESSENTIALISM TO INDIVIDUALISM

Chemical elements and biological species have been the two stock examples of natural kinds from Aristotle to Putnam. Samples of lead or individuals of the species *Pan troglodytes* are not only similar in various respects, but are “of the same kind” in some much deeper sense. One way to express this deeper commonality is to say that the members of a kind share an essence—a property common to all the members of a kind and responsible for each member being the kind of thing that it is. My gold watch resembles your gold navel ring in many respects, some known to us and some not, because the atoms of which both are composed share an essence: their atomic number. Essentialism in biology would suggest that my kitten Erasmus resembles Socks the cat because they too share an essence (albeit a less well understood one). Essentialism took on a new life in the 1970s, largely because of the work of the philosophers Saul Kripke (1980) and Hilary Putnam (1975). Biological species were one of the stock examples in this essentialist literature, even though by this time essentialism was regarded by many biologists as inconsistent with the basic tenets of Darwinism!

The perceived antithesis between evolution and essentialism was largely due to the work of Ernst Mayr (1959). Mayr argued that biology before Darwin was characterized by *typological thinking* in which types or kinds of organisms had ontological and explanatory priority over concrete individuals. Darwinian *population thinking* gives populations of concrete individuals ontological and explanatory priority instead. Elliott Sober has argued convincingly that the core of population thinking is the Darwinian approach to variation (Sober 1980). The typological approach explains the resemblances between the individuals in a species in terms of the underlying “natural state” of each individual, just as chemistry explains the resemblances between atoms of the same element in terms of their shared microstructure. The typological approach explains variations between the individuals in a species as perturbations of the natural state of that species. The Darwinian approach explains both resemblance and variation at the population level. Organisms resemble

one another not because of something inside each of them, but because of something outside each of them: the genealogical and ecological factors that make these organisms a population or a group of related populations. The properties that differ between individuals are ontologically on a par with those properties they share. Variation is not noise obscuring the essential sameness of the members of a species, but an important, heritable property of populations consisting of the aggregate real differences between its members. Sober concludes that because these explanations of sameness and difference are central to the Darwinian tradition, Mayr is correct in concluding that Darwinism precludes identifying any phenotypic or genotypic features as a species essence. However, Sober notes that it would be quite consistent to be a Darwinian essentialist, given the right choice of essential properties (1980, 209). Population thinking excludes essential intrinsic properties, but it does not exclude essential relational properties. This paper defends just such a relational essentialism.

Mayr famously tried to characterize the relational properties that unite the members of a species. His *biological species concept* (BSC) defined species as “groups of actually or potentially interbreeding natural populations which are reproductively isolated from other such groups” (Mayr 1940, cited in Mayr 1963, 19). Underlying this formal definition of species in terms of interbreeding is the idea of a genealogical nexus. A nonreproductive worker in one beehive neither actually nor potentially interbreeds with nonreproductive workers in other hives, but that individual is united in a genealogical nexus with reproductives who actually or potentially interbreed with reproductives in other hives. These reproductives in turn are united in a genealogical nexus with nonreproductives in their hives, so the several nonreproductives are members of the same species. Attempts to extend the BSC to asexual species also rely on this underlying genealogical element in Mayr’s species concept, using it as one of two criteria for specieshood. According to such proposals, an asexual species is a well-defined segment of a genealogical tree of asexual individuals that meets some other criteria, such as containing individuals roughly as morphologically similar to one another as members of a sexual species. This second criteria is designed to distinguish the species-level genealogical tree segment from the larger segments in which it is embedded and from the smaller segments that it embeds. Robert Brandon and Brent Mishler (1987) have generalized these two criteria for specieshood into *grouping criteria* and *ranking criteria*, and they have argued that any species concept must have both a grouping and a ranking criteria. In most modern species concepts, including modern versions of the BSC, the grouping criteria is genealogical. Species must be characterized by some version of monophyly—descent from a single population, a single speciation event, or any similar unique point of origin. The ranking criteria serves to distinguish species from equally monophyletic genera, families, and so forth. Although there is some disagreement over the best definition of species-level monophyly, the main disagreements between the twenty or so current species concepts are in their

different ranking criteria. (For a very clear look at species-level monophyly, see Kornet 1993.)

*Individualism* about species is an idea with close links to antiessentialism, both conceptually and historically. Individualists argue that species are not kinds or types at all, but big individual objects. Organisms are not *members* of a species, but *parts* of a species. The individualist arguments of Michael Ghiselin (1974a, 1974b) and David L. Hull (1976, 1978) strongly resembled some of the earlier arguments against essentialism. They argued that species must be able to evolve and that kinds or types do not evolve. Species must be able to undergo unlimited change in any of their genetic or phenotypic characters, not only in peripheral characters. If species were kinds or types of organism, then it would not be species that evolved, but organisms or populations that changed from one species into another. The kinds or types would form a sort of biological absolute space against which evolutionary change occurred. Hull and Ghiselin also pointed to the practical failure of attempts to define species by lists of characters or statistical clusters of characters. This argument took on new force in light of the cladistic revolution in systematics, in which attempts to discern common descent replaced attempts to identify taxa by statistical clusters of characters. These and other arguments convinced the individualists that species could be defined only in terms of the pattern of ancestry and descent among organisms, but the next step in their reasoning is the most relevant to the issues of this paper. Hull and Ghiselin concluded that because species and other taxa must be defined in terms of genealogy, they must be moved from the ontological category of types or kinds to the category of individual objects. If taxa are genealogically or historically defined, then they cannot be natural kinds.

This last step in Ghiselin and Hull's argument depends on a traditional conception of natural kinds in which they are the subjects of spatiotemporally unrestricted laws of nature. If natural kinds are to figure as the subject of universal laws, they themselves must have universal applicability. Laws that make ineliminable mention of things that can exist only at a particular location in time and space are not, in the relevant sense, universal laws. If Ghiselin and Hull are correct, then biological taxa have just such a unique origin in space and time. No part of a taxon can exist outside the cone of causal influence extending from its origin event, so taxa are restricted to a particular portion of space-time and cannot be mentioned in genuine laws of nature.

The conclusion that there are no laws of nature concerning taxa has been welcomed by many theorists as part and parcel of antiessentialism. Hull (1986) has welcomed the liberatory conclusion that there is no such thing as "human nature." Attempts to distinguish normal from abnormal humans are simply misguided. John Morss (1992) has argued that there are no laws of ontogeny and particularly of child development. We should be suspicious of theories that describe a series of stages through which every child passes to reach maturity. The downside of the anomalousness of biological taxa is that

it threatens the status of biology as a science. If there are no biological laws, biology is merely the study of how things happen to be around here right now (Smart 1963). The threat is not merely to laws about species and speciation. The parts of organisms and their physiological processes are standardly classified in the same way as whole organisms—using the Darwinian or evolutionary homology concept: two organs are homologous if they are copies of a single ancestral organ. Thus, the wings of pigeons are “the same” as the wings of albatrosses in a way that they are not “the same” as the wings of fruit bats. The conclusion that there are no lawlike principles of the structure and development of organs or physiological processes has not been welcomed by developmental and structural biologists.

Hull’s response to these worries is interesting and has not been sufficiently discussed. He notes that there are two quite different schemes of classification in biology. Systematics and perhaps the anatomical and physiological disciplines classify by homology. If this scheme is evolutionary homology, they face the problem we have just encountered. Ecology and functional biology, however, seem to classify by analogy. Kinds such as predator, prey, digestion, thermoregulation, and so forth are not genealogically defined. Entirely unrelated organisms can share an ecological role. The wings of pigeon and fruit bat may be “the same” in the sense that they are both shaped for work amongst the branches. Likewise, genealogically unrelated DNA sequences can code for a protein with the same metabolic function. Hull suggests that it is to these categories of analogy that biology must turn in its search for laws. This suggestion is attractive when put in these abstract terms, but when we try to apply it, the results are worrying in the extreme. Developmental and structural biologists classify by homology for good reason. Functional resemblances between organs tend to be shallow. In human engineering, devices that have the same function but that were designed independently tend to be very different. In the same way, the circulatory system of an octopus is very different from that of an aquatic mammal of similar size. If developmental biology and structural biology seek only laws about functional kinds, then laws in these disciplines may be little more than performance specifications (Griffiths 1994, 1996a, 1996b). Hull’s recommendation also came just as ecologists were turning from the dreams of a grand theory that had occupied them in the 1960s to a renewed interest in contingency and history (Kingsland 1985). Ecological models, it was suggested, may never achieve the status of universal laws and will remain always in need of testing and retuning for each new case. This trend in ecological thought has continued, so if biology looks to ecology for its spatiotemporally unrestricted laws of nature, it may not be pleased with what it finds.

### **DOES BIOLOGY NEED NATURAL KINDS?**

A number of biologists have argued that biology cannot do without natural kinds. The *process structuralist* school has suggested that biology has no real

explanation of form unless it has an explanation in terms of natural kinds (Goodwin and Saunders 1989, Goodwin 1994). I have argued against this claim elsewhere (Griffiths 1996a), but it has recently been reiterated (Webster and Goodwin 1996). There is much that is correct in this new presentation, but I argue that what is correct can be accommodated by a thoroughly Darwinian and historical conception of biological kinds. Other authors have argued that understanding the nature of the characters that are candidates for evolutionary homology will require a nonevolutionary, structural concept of character identity (Müller and Wagner 1991; Wagner 1994, 1996). I consider the relations between the evolutionary homology concept and this proposed structural-developmental homology concept in my closing section.

In their 1996 book *Form and Transformation*, Brian Goodwin and Gerry Webster reiterate their claim that biology needs natural kinds of organisms, parts, and processes. If experimental biologists are to perform repeatable experiments, they must be able to say what it would be to have subjects "of the same type." If the anatomical and physiological disciplines are to extrapolate from the individuals in the laboratory to individuals elsewhere, they need to know what *sort* of things they have been examining. Goodwin and Webster argue that scientific practice only makes sense on the assumption that there are real *sorts* of things as well as real *individual* things. Individuals of the same sort share some underlying "nature," and it is the aim of science to elucidate these "natures." Goodwin and Webster agree with the anti-essentialists that the shared nature of a biological kind cannot be either a set of phenotypic characters or a set of genes. Both vary too much within the groups, such as species, about which biologists need to generalize. They infer from this variation that organisms of the same sort must have something else in common: something that must emerge as essentially the same in each individual despite differences in the developmental resources that individuals inherit. It must also be something that can be shared by many phenotypically different individuals. Differences within a species must be varied expressions of a common underlying nature. Goodwin and Webster thus reject the population-thinking model of variation in which the properties that differ between individuals are ontologically on a par with those properties that are shared. They revert to a typological model, in which variation is explained as the response of an essentially similar system to different inputs. A genetic change or a change in some other developmental input acts on the "real nature" of the organism to cause it to express a new outcome among the range of outcomes that it is, in Webster and Goodwin's phrase, "competent" to produce. Variation masks the real underlying sameness of a type of organism and it is the task of science to see through the variation to the essential sameness.

Goodwin and Webster's candidates for the real natures of organisms, parts, and processes are *morphogenetic fields*. They conceive of these fields as an emergent level of organization in the developmental process. The existence of such a level of organization explains the constancy of biological

form in the face of substantial variation in all elements of the developmental matrix, including the genome. Goodwin is fond of comparing the morphogenetic field to an attractor in complex systems theory. Development from a wide range of genetic starting parameters is drawn to an attractor represented by a particular morphogenetic field (Goodwin et al. 1993). The existence of such an emergent level of organization can also explain the fact that mutations and phenocopies are often equivalent. The abnormal bithorax phenotype in *Drosophila* can be produced by a genetic change (the bithorax mutant) or by an environmental change (the bithorax phenocopy). The potential to produce the bithorax form is thus inherent in the morphogenetic field of the segment that becomes a second thorax. This potential can be triggered by several different perturbations to that field.

A revival of the morphogenetic field concept has also been advocated recently by Scott Gilbert, John Opitz, and Rudolph Raff (1996). Their conception of a morphogenetic field is much closer to the idea of a gene control circuit: a set of genes linked together by relations of feedback, excitation, and inhibition. Their field concept is directed at explaining the same phenomena as Goodwin and Webster's. By linking many genes together in this way, they hope to explain the sense in which development is an emergent phenomena: a circuit may have properties that are robust when some constituents of the circuit are changed and may be pushed into the same alternative configuration by any of several different perturbations. Despite these similarities, there is a critical difference between the two field concepts. Webster and Goodwin strongly resist the idea that a field can be reduced to the genes and other molecular machinery that underlie it. They also resist identifying the "competence" of the field with the norm of variation of those genes, mainly because their field is essentially an *invariant* across individuals of the same kind. Changing the particular genes that underlie the field *makes no difference* to the field itself. When a genetic change causes a phenotypic change, according to Goodwin and Webster, we are not seeing the result of a slightly different morphogenetic field, but an identical field producing another of the outcomes within its competence. Gilbert, Opitz, and Raff make no such essentialist commitment, which demonstrates an important conceptual point. It is not necessary to postulate a theoretical entity to act as an absolute invariant in order to explain robust developmental outcomes. The field concept is entirely viable in a population-thinking form in which robust developmental outcomes are explained by the fact that many *different* (but similar) morphogenetic fields produce the same outcome.

Goodwin and Webster's case for the existence of morphological fields as developmental invariants is driven not by the need for an emergent level of developmental organization to explain canalization and mutation/phenocopy equivalences, but by the abstract methodological claim made at the beginning of this section. They do not see how extrapolation from observed to unobserved instances can be valid unless these instances share some underlying invariant nature. They postulate invariant morphogenetic fields to

meet this *epistemological* need. In the next two sections, I explain why this postulation is a misunderstanding of the natural kind concept and of how natural kinds really license such extrapolation.

## NATURAL KINDS WITHOUT TEARS

Induction and explanation presume that the world contains correlations between properties that are, to use Nelson Goodman's term, *projectable* (Goodman 1954). We can depend on these correlations holding in new cases. Theoretical categories embody current understanding of where such projectable clusters of properties are to be found. The species category, for instance, is supposed to reliably collect morphological, physiological, and behavioral properties. We can investigate these properties in the species as a whole by studying a few members of the species. That being accomplished, we can explain the fact that an individual has certain properties by citing its species: any organism that *was* of this species *would* have those properties. In Goodman's original presentation, the projectability of theoretical categories is supposed to be judged on the basis of our past experience in using the categories and others related to them—which, in practice, means that we judge projectability on the basis of our background theories of the domain to which a theoretical category applies. Our theories lead us to believe that all the chemical properties of sulphur will be reliably reproduced by future instances of that element, whereas few if any of the physical properties of Citroen cars will be reliably reproduced by future instances of that marque. Natural kinds are simply a realist interpretation of Goodman's projectable categories. The categories that figure in successful theories are projectable because the theories have some degree of *versimilitude*. The instances of these categories really do share an underlying nature. Therefore, from the realist perspective they adopt, Webster and Goodwin are correct in claiming that for biology to engage in induction and explanation, it must have theoretical categories that represent natural kinds.

The concept of a natural kind has a long history, stretching back at least as far as John Locke's discussion of the distinction between *real kinds* and *nominal kinds*, if not as far as Plato's famous remark about "carving nature at its joints" (Hacking 1991a). In the logical empiricist tradition, from which philosophy of science as we now know it emerged, natural kinds are envisaged as the objects of spatiotemporally unrestricted laws of nature. They are the nodes around which theories in the fundamental sciences are structured (Quine 1977). But recent decades have seen substantial changes in thought about natural kinds. Greater philosophical attention to the special sciences has led to the eclipse of the idea that these sciences are one day to be reduced to more foundational sciences such as physics and chemistry. The current received view is that the dynamics of physical systems can only be adequately captured using a hierarchy of theoretical vocabularies, each irreducible to the vocabularies below it. Irreducibility is guaranteed by the fact

that descriptions in one vocabulary can be made true by indefinitely many arrangements of the structures described in lower-level vocabularies (Fodor 1974; Wimsatt 1976a, 1976b; Jackson and Pettit 1988; Lycan 1990). There are indefinitely many ways, for example, to construct instances of money, a central theoretical category of economics. An empirically successful theory with such irreducible categories cannot be eliminated without losing the knowledge embodied in its empirical generalizations. Economic generalizations about money, for example, can be made true by indefinitely many physical systems of currency and so cannot be replaced by generalizations about any category of physical systems. This idea has led to what Richard Boyd (1991) has called "the enthusiasm for natural kinds" (p. 127). Categories from any special science that enter into the generalizations of that science are now commonly regarded as natural kinds. Inflation and schizophrenia take their place alongside electrons and stars.

The generalizations of the special sciences often fail to live up to the ideal of a universal, exceptionless law of nature. Generalizations in psychology or economics are often exception-ridden or hedged with generous *ceteris paribus* clauses or both: decreases in the money supply *usually* lead to a contraction of the economy, *all other things being equal*. Nevertheless, the key feature of a law of nature is still present in these generalizations: they have *counterfactual force*. The idea of counterfactual force is central to the traditional idea of a law of nature because it explains how laws differ from mere widespread coincidences. It may well be true, for example, that every species with an eusocial grade of social organization has individuals that weigh less than 5,000 kg, but even if this statement turns out to be true throughout the Federation of Planets, it will not be a law of nature. Nothing in our theories licences the subjunctive conditional, "if this *were* a member of an eusocial species, it *would* weigh less than 5,000 kg." This statement lacks counterfactual force: it is not "lawlike." A key part of the conception of a natural kind is that it is a category about which there are lawlike, counterfactual-supporting generalizations. We can use induction to investigate natural kinds because we expect certain classes of properties to be connected to those kinds in a lawlike, rather than a coincidental way. For example, our background theories licence the expectation that samples of an element will possess their chemical properties in a lawlike rather than a coincidental way. Having tested the chemical properties of the samples, we can extrapolate to the chemical properties of other instances of the element.

The idea of counterfactual force is easily generalized to the exception-ridden generalizations of the special sciences. Minimally, any generalization that is a better predictor of phenomena than a suitably designed null hypothesis has some counterfactual force. This allows us to frame a minimal conception of *naturalness* for kinds. A kind is (minimally) natural if it is possible to make better than chance predictions about the properties of its instances. Surprisingly, this utterly minimal conception of a natural kind is not

toothless. It does not license the conclusion that any way of classifying nature is as good as any other. Natural kinds are ways of classifying the world that correspond to some structure inherent in the subject matter being classified. They contrast to arbitrary schemes of classification about which the nominalist claim that the members of a kind share only a name is actually true. Furthermore, the minimal account of naturalness lends itself to successive restrictions that allow us to distinguish between kinds of greater or lesser naturalness and hence of greater or lesser theoretical value.

Although it is not possible in this essay to give an adequate treatment of the principles for choosing between alternative taxonomies of nature, a brief, general outline may be helpful. The value of a lawlike generalization can vary along two independent dimensions, which we might call *scope* and *force*. *Force* is a measure of the reliability of predictions made using that generalization. *Scope* is a measure of the size of the domain over which the generalization is applicable. A theoretical category about which there are generalizations of considerable scope and force is more natural than one about which generalizations tend to have more restricted scope and lesser force. For example, the claim that cladistic taxonomy is “maximally predictive” of the unobserved properties of taxa is intended to show that cladistics is superior to other systems in terms of force. There will not always be a clear winner when we compare two sets of theoretical categories on the basis of scope and force. Scope and force may trade off against one another. The scope of generalizations made with one set of categories may overlap rather than include the scope of generalizations made with the other taxonomy so that neither taxonomy can be discarded without loss of understanding.

Theoretical categories can also differ in the number of generalizations into which they enter so that one category can seem the focus of a richer scientific project than another, irrespective of comparisons of the strength of the generalizations they yield. Finally, theoretical categories are tied up in wider research programs whose relative prospects may cause us to prefer that set of categories to another despite a paucity of currently established generalizations about the preferred set of categories. None of these considerations, however, refutes the basic idea that some theoretical categories are superior to others and that some are of no foreseeable value whatever. Even if different categories are valuable for different purposes, it is still true that some are better for a particular purpose than others and that some have no foreseeable use at all. Embodying these ideas in the language of natural kinds links it to a broadly realist perspective in which the predictive and explanatory value of categories is taken to be *prima facie* evidence that they capture part of the structure of the world. The “enthusiasm for natural kinds” embodies the realization that there is more structure in the world than can be captured by a single taxonomy of nature.

Richard Boyd has outlined a similar conception of natural kinds using his idea of *causal homeostasis* (Boyd 1991). According to Boyd, we judge a kind

to be projectable, or natural, when we have theoretical grounds for supposing (or we simply postulate) that there is a causal explanation for the property correlations we have observed. Boyd calls this underlying reason a *causal homeostatic mechanism*—something that causally explains the maintenance of the same property correlations throughout the set of instances of the kind. In my reading of Boyd, this causal homeostatic mechanism corresponds to the traditional “essence” of a natural kind. In the paradigmatic example of chemical elements, the causal homeostatic mechanism is a shared microstructure. It is because of their subatomic composition that the instances of a chemical element share their chemical properties. However, nothing in the idea of a causal homeostatic mechanism requires the mechanism to take the form of a set of intrinsic properties possessed by every member of the kind and synchronically causally producing the other properties characteristic of the kind. Money, for example, has no such microstructural essence, although it is a key node in many economic theories. The lawlike generalizations about money, such as those connecting money supply to inflation or to interest rates, hold true in an economy because of a social convention treating some class of objects as a means of exchange and because agents in that economy try to maximize their utility. Neither of these circumstances is linked to any intrinsic property of the currency units. In a similar way, if characteristic ecological successions represent natural kinds in ecology, the causal homeostatic mechanism for the kind “Fiordland rainforest succession” will include the available range of seeds and other propagules, the climate of the region, and so forth. All that is required for the existence of a natural kind is that there be some causal process in nature that links together several different properties of the objects influenced by that process. A shared microstructure is only one way of achieving this “homeostasis” of properties.

The idea of a causal homeostatic mechanism frees the idea of *essence* from many of its traditional commitments—commitments that have proved problematic in the case of biology. My interpretation of Boyd’s work is that he provides a general analysis of the role that essences play in scientific reasoning about natural kinds and then redefines essence as any property that can play this role. Any state of affairs that licenses induction and explanation within a theoretical category is functioning as the essence of that category. The essential property that makes particular instances members of the kind is their relation to that causal mechanism, whatever it may be. One exciting implication of this approach is that it breaks down the traditional distinction between natural kinds and kinds generated by human agency. I have exemplified this possibility by using money as an example of a natural kind. Artifactual kinds, such as kinds of tool or ceremony, can be the subject of lawlike generalizations because the sociological causes that produce them can function as essences. These sociological causes guarantee with some degree of reliability in some suitably delimited domain that instances of the kind will share a cluster of properties.

Boyd's proposal is a substantial revision of the traditional ideas of essence and natural kindhood (see also Boyd, chapter 6 in this volume). "Natural" kinds that have never been seen before can be created by social processes unique to a particular society. The fact that people think certain things form a kind can function as the essence of that kind! The justification for these conceptual revisions is that they allows insights about the formation and use of theoretical categories to be extended to the special sciences rather than restricted to a (dwindling) core of kinds with microstructural essences. The psychologist Frank C. Keil (1989) has used Boyd's ideas to argue for a continuity between category formation by developing children and category formation in science (see also Keil and Richardson, chapter 10 in this volume). I have argued that the formation of theoretical categories in psychology, including categories unique to particular cultures, is best understood as a search for causal homeostasis (Griffiths 1997).

In this section, I have tried to motivate a very general conception of natural kinds, one that discards many of the traditional associations of the natural kind concept. Natural kinds are needed for induction and explanation. They represent theoretical categories that we judge to be projectable, which requires them to enter into lawlike, counterfactual supporting generalizations. It does not require that these generalizations be universal, deterministic laws: lawlike generalizations of more limited scope and force are enough. Finally, kinds are defined by the processes that generate their instances, and for many domains of objects, these processes are extrinsic rather than intrinsic to the instances of the kind. The causal homeostatic mechanism that guarantees the projectability of a kind plays the traditional role of an essence, but it need not be a traditional, microstructural essence.

## HISTORICAL ESSENCES

Cladistic taxa and parts and processes defined by evolutionary homology have historical essences. Nothing that does not share the historical origin of the kind can be a member of the kind. Although Lilita might not have been a domestic cat,<sup>1</sup> as a domestic cat she is necessarily a member of the genealogical nexus between the speciation event in which that taxon originated and the speciation or extinction event at which it will cease to exist. It is not possible to be a domestic cat without being in that genealogical nexus. Furthermore, cladistic taxa and parts and processes defined by evolutionary homology have no other essential properties, which is why process structuralists such as Goodwin and Webster do not think that these categories can be adequate for developmental and structural biology. They do not see why kinds whose only essential properties are historical should be the subjects of lawlike, counterfactual-supporting generalizations about morphological and physiological properties. Yet there is a well-known Darwinian ground for expecting groups defined by common descent to share morphological and physiological characters:

It is generally acknowledged that all organic beings have been formed on two great laws—Unity of Type and the Conditions of Existence. By unity of type is meant that fundamental agreement in structure, which we see in organic beings of the same class, and which is quite independent of their habits of life. On my theory, unity of type is explained by unity of descent. (Darwin 1859, 206)

Even in its most extreme adaptationist forms, Darwinism retains these two “great laws” as separate forces that conjointly explain biological form. The principle of heredity acts as a sort of inertial force, maintaining organisms in their existing form until some adaptive force acts to change that form. This *phylogenetic inertia* is what licenses induction and explanation of a wide range of properties—morphological, physiological, and behavioral—using kinds defined purely by common ancestry. If we observe a property in an organism, we are more likely to see it again in related organisms than in unrelated organisms. Since Darwin, this idea, much elaborated, has been the basis of comparative biology (Brooks and McLennan 1991, Harvey and Pagel 1991).

However, the mere existence of phylogenetic inertia is not the whole story. There are striking contrasts between biological traits in their tendency to persist without reference to the “conditions of life.” I have argued elsewhere that it is a mistake to assume that when we have a selective explanation for the origin and fixation of a trait, there is nothing left for selection to explain (Griffiths 1992, 1996b). Many traits display a pattern of phylogenetic inertia reminiscent of the inertia of Aristotelian physics. Just as early physics expected a body with no forces acting on it to return to rest, these traits tend to atrophy when no selective forces work to maintain them. The apparently panphyletic tendency of cave-dwelling organisms to lose pigmentation and sight is a well-known example. With traits displaying this pattern, selective explanations of their maintenance are as legitimate as selective explanations of their origin. In contrast to these Aristotelian traits, other traits display an apparently Newtonian pattern of phylogenetic inertia. They are maintained over the longest geological timescales and the widest range of conditions of life, with no apparent regard for adaptive utility. Traits of this kind are the ones that make good taxonomic characters. The pattern of fused segments that marks out crustaceans among the arthropods is a well-known example, and classic morphological traits like this are not the only sort of traits that display the Newtonian pattern. Part of Konrad Lorenz’s legacy was the realization that some behaviors also have a very strong phylogenetic signature.

The fact that different traits display such different patterns of phylogenetic inertia calls out for a developmental explanation. Development is the obvious place to look for something that reduces variance in certain traits and so causes them to resist atrophy or elimination as an effect of adaptive change. Proposals for developmental explanations of strong phylogenetic inertia can be divided into two types. The first type includes Rupert Riedl’s (1977) concept of *burden* and William C. Wimsatt’s (1986) notion of *generative entrench-*

ment (see also Schank and Wimsatt 1986). Both concepts draw attention to the fact that one trait may be developmentally linked to a range of other traits, making its elimination far less likely than if it were an independent developmental unit. Proposals of this type still assign a major role to selection in maintaining traits, although it acts indirectly through the structure of the development system. The second type of developmental explanation of phylogenetic inertia avoids implicating selection in any way. Examples of this type include Goodwin and Webster's concept of *generic forms* and perhaps some of Stuart Kauffman's ideas (Goodwin et al. 1993, Kauffman 1993, Webster and Goodwin 1996). These approaches take the fact that a trait is widespread in a group as a sign that this trait is an easy one for that kind of developmental system to generate. The widespread occurrence of the trait is not to be explained by its utility or its links to other useful traits, but by the structure of the developmental system. In one of Goodwin's favorite examples, the fact that there are only three patterns of phylotaxis in higher plants—patterns in which successive leaves emerge from the stem—is explained by the existence of three stable attractors that emerge when a single, continuous, quantitative developmental parameter is altered in a model of the growth of the meristem. The spiral phylotactic pattern seen in 80% of these plants emerges from the model as the outcome with the largest basin of attraction (Goodwin 1994, 116–133).

Both types of explanation of phylogenetic inertia support rather than oppose the idea that categories based on evolutionary homology will provide a natural taxonomy with which to investigate morphological and physiological characters. Past discussion of the second type of explanation has tended to give the opposite impression, however. The reasons for this tendency lie in philosophy rather than biology. Goodwin, Webster, and other process structuralists have argued that categories based on evolutionary homology do not have an underlying "nature" suitable for scientific investigation because evolutionary homologies do not have traditional, microstructural essences. They infer from this argument that if biology is to be scientific, biological kinds with such essences must exist. Furthermore, because there are some reliable taxonomic characters—the Newtonian traits just discussed—they infer that these characters must have just such underlying microstructural essences. I tried to show in the last section that the philosophical part of this process structuralist argument is mistaken. Microstructural essences are not needed to justify explanation and induction. What is left of the argument is just the postulation of developmental causes for phylogenetic inertia, with which the Darwinian can wholeheartedly agree. However, because of her general theoretical orientation, the Darwinian will have expectations very different from the process structuralist's about these developmental mechanisms. She will expect them (*a*) to have a phylogenetic pattern like other characters, and (*b*) to show variation in natural populations. The first of these expectations supports the continued use of historically defined kinds in biology, including biological investigations of the

developmental basis of phylogenetic inertia. The second expectation means that even when the developmental basis of phylogenetic inertia is understood, the Darwinian will not expect to see historical kinds displaced by purely developmental definitions of taxa, parts, and processes. In the next two sections, I expand on these two points.

### WHY GHISELIN AND HULL WERE WRONG

Antiessentialists and individualists about biological taxa were wrong to suppose that there are no lawlike generalizations about these taxa. A hierarchical taxonomy based on strict phylogenetic principles will collect more of the correlations between characters, from molecular to behavioral, than any other taxonomy we know how to construct. Such a taxonomy will group organisms into natural kinds because it will predict with considerable force many properties of individuals. Although such a taxonomy will predict the properties of unobserved genera or species, it will function most powerfully in predicting the properties of new members of taxa at or below the species level. A number of competing (though not necessarily exclusive) explanations of the special status of species are embodied in some of the twenty or so currently proposed species concepts. These explanations draw attention to causal processes such as gene exchange (biological species concept) or selection for the requirements of a niche (ecological species concept). These mechanisms reinforce phylogenetic inertia in keeping the members of a species clustered together in the space of biological possibility (cf. de Queiroz, chapter 3 in this volume).

Generalizations about taxa are exception-ridden. This does not, however, prevent them from being lawlike or having counterfactual force. The causal homeostatic mechanisms of taxa license the prediction that a new bird will detect its prey using visual cues or that in a new cephalopod, the blood vessels supplying the retina will lie under rather than over it. The causal homeostatic mechanisms also make it legitimate to extrapolate experimental results to other members of the same taxon, especially at the species level. The fact that such predictions and extrapolations are not absolutely reliable is simply beside the point. They are more reliable than chance, so unless there is some other way to capture the same regularities, eschewing the use of these categories would mean discarding some of our understanding of the structure of nature.

Parts and processes defined by evolutionary homology can be used for explanation and induction for the same reason that historically defined taxa can be used: phylogenetic inertia licenses the extrapolation of morphological and physiological properties in categories defined by common ancestry. Also, among these properties are the very developmental processes that are likely to explain the phenomena of phylogenetic inertia. Developmental processes, as much as other anatomical or physiological kinds, can be expected to reflect phylogeny. What lies at the bottom of all these phyloge-

netic patterns is, after all, the fact that related organisms inherit similar developmental resources. Plant physiology, for example, does not converge on animal physiology whenever it would be adaptively useful for it to do so because plant cells inherit a range of membrane templates, organelles, genes, and so forth that are fundamentally different from those inherited by animal cells.

If historical kinds are natural because related individuals inherit similar developmental resources, it might seem possible to define the kinds in terms of the developmental resources that underly them. This proposal would treat the shared developmental resources of a taxon as the causal homeostatic mechanism of that taxon, a mechanism that takes the form of a traditional microstructural essence possessed by all and only the members of the taxon. I suspect that this thought is at the back of several structuralist criticisms of the use of historical kinds in biology. The structural or developmental biologist sees that the processes they are investigating explain the fact that members of a taxon share a rich cluster of properties, which suggests to them that the real essence of the taxon is not its shared history, but its shared developmental processes. In the next section, I show that this very natural line of reasoning is mistaken because of the original, Darwinian considerations against essentialism outlined at the beginning of the paper.

#### WHY GHISELIN AND HULL WERE RIGHT

Darwinians will resist the suggestion that taxa be defined developmentally because they expect developmental processes to be just one more product of evolution. As such, they expect developmental processes to display variation between individuals in natural populations, just as other characters do. Empirically, they do not expect to find a list of developmental properties possessed by all and only the members of a species any more than they expect to find lists of phenotypic or genotypic characters possessed by all and only members of a species. Conceptually, even if such a list of properties existed for a species, it would be an accidental not an essential matter. An individual united in a genealogical nexus with the existing members of the species, but lacking some property on the list, would still function as a member of the same evolutionary unit. The purpose of the species concept for a Darwinian is to describe the units of evolution, and essentialist species concepts fail to do this.

There are a number of reasons why Darwinians have wanted to take a more developmental perspective on evolution. A developmental perspective highlights the problems with an atomistic approach to the evolution of characters, in which each character is assumed to be optimized independently of the others (Gould and Lewontin 1979, Lewontin 1983). A developmental perspective can also draw attention to the wide range of developmental resources other than genes that can be the subject of evolutionary explanation (Griffiths and Gray 1994). Perhaps most importantly, a developmental

perspective may allow a more adequate integration of phenotypic and genotypic evolution. This last motivation is at the heart of Gilbert, Opitz, and Raff's (1996) proposal to revive the morphogenetic field concept. They suggest a definition of evolution as change over time in the developmental biology of a lineage, which contrasts with the currently popular definition of evolution as change in gene frequencies in a lineage. But all these goals of a *Darwinian developmentalism* require development to be part of the process of evolution by natural selection. As such, development must be something that exhibits heritable variation. It cannot be something that is invariant across all the members of a species.

The Darwinian developmentalist is an evolutionist who focuses on development, just as a gene selectionist is an evolutionist who focuses on genes. These two views of the evolutionary process differ in important ways, but they agree on some central Darwinian themes. It is these themes that Ghiselin and Hull were right about in their insistence on a historical, anti-essentialist view of taxa and of homology. One central Darwinian theme is the ubiquity of variation. Where the Darwinian developmentalist observes a widespread phenotypic character, she will not assume that it is produced by an underlying, developmental invariant. She will be open to the idea that it is an outcome that can be produced by any of a range of different but similar developmental processes. *Canalized* developmental outcomes are precisely those that can be produced by many different configurations of developmental resources. Developmental biology illuminates how canalization occurs, but it need not do so by finding or postulating a developmental invariant other than the canalized outcome itself. Another central Darwinian theme is the value of a phylogenetic perspective in all biological investigations. The Darwinian developmentalist will expect to find a phylogenetic signature in characters of all kinds and to make extensive use of the comparative method in testing hypotheses about character associations. This phylogenetic perspective will extend to developmental biology.

### THE STRUCTURAL HOMOMOLOGY CONCEPT

I have defended the view that historically defined taxa are natural kinds and the corollary view that evolutionary homologues are also natural kinds. I have defended these views against some arguments associated with structuralist approaches to biology. In this closing section, I want to consider two other, recent arguments that biology needs a structural homology concept. The first argument suggests that the evolutionary homology concept is somehow unworkable without a prior conception of structural homology. This argument is mistaken, but a second, better argument points to the potential value of a structural homology concept, including its value in illuminating the basis of the evolutionary homology concept.

The mistaken argument, which we can perhaps regard as put forward by a hypothetical structuralist strawman, is that because candidates for evolu-

tionary homology must be real characters of organisms, the identification of evolutionary homologues is parasitic on the identification of characters defined by some nonevolutionary homology concept. It is certainly true that before it can be asked whether two characters in different taxa are homologous or homoplastic, they must be identified as characters. We might, for example, measure the ratio of length to circumference of a bone, find that it was constant across a range of taxa, and use a cladistic analysis of a whole suite of characters to determine if this commonality can plausibly be identified as a homology. The first part of this procedure embodies a decision to treat the ratio as a character. It is also true that not everything that can be measured is a real character. Probably no one would bother to measure in different taxa the ratio between number of retinal receptor types and number of legs. However, it does not follow that we need to know which features of organisms are real characters before we start looking for homologies. Cladistic analysis can proceed from a list of arbitrary measurements by looking for congruences among the evolutionary trees produced by different measurements and thus "bootstrapping" itself into a reliable character set. A set of characters, different subsets of which produce similar trees, is probably a set of real units of inheritance and evolution.

The better argument for the desirability of a structural homology concept is given by Gunther P. Wagner (1994; see also Müller and Wagner 1991, Wagner 1996). The pre-Darwinian homology concept distinguished homologous resemblances among taxa from analogous ones. *Homologies* are different instances of the very same character, whereas *analogies* are different characters that happen to resemble one another. Darwin gave a specific interpretation to this idea of being *really* the same character rather than *apparently* the same character. Two characters are really the same if they are both the same as some character possessed by a common ancestor. Wagner's point is that this definition does nothing to explain the sense in which characters are "the same" by descent. Darwin has analyzed character identity horizontally, between taxa, but not vertically, between parent and offspring. It is simply assumed that some resemblances between parent and offspring amount to true character identity, just as it was previously assumed that some resemblances between taxa amount to true character identity. Wagner's point is not that the Darwinian needs to understand the vertical relation of character identity before she can begin to reconstruct phylogeny. As I have argued, the Darwinian can simply presume that there are real units of inheritance and identify good candidates for these units by trial and error. The point, rather, is that until we understand the nature and origins of the units of heritable biological form, we will not know *why* this bootstrapping procedure works. More generally, we will not understand why the historical, phylogenetic approach to biology is so useful. To explain this fundamental fact about biology we need to understand why some characters and not others display phylogenetic inertia. As Wagner (1994) puts it, "the main goal of a biological [i.e., developmental] homology concept is to explain why certain parts of

the body are passed on from generation to generation for millions of years as coherent units of evolutionary change" (p. 279).

It is this sort of question that has been the focus of Wagner's more recent work on the evolution of modularization and canalization of development (Wagner 1996, Wagner et al. 1997). Wagner (1994) rejects an analysis of vertical character identity based on identical developmental origin, a modern derivative of the traditional practice of judging homology from the relative position of parts in the embryo: "Too often do we find substantial developmental variation among structurally, and presumably phylogenetically, identical body parts" (p. 276). He would presumably reject an account of vertical character identity based on identical genetic causes for the same reason: homologous characters can persist through substantial changes in the genetic inputs to their development. In place of such ideas, Wagner sets up the goal of understanding why organisms have come to have discrete, re-identifiable parts. A theory of why there are parts will tell us how those parts can be naturally taxonomized. Wagner's research program is thus (quite self-consciously) a search for natural kinds construed as the objects of lawlike generalizations.

Wagner's work is an instance of what I have described as "Darwinian developmentalism" because he looks for the origins of these units of structural homology in the evolutionary process rather than in a system of ahistorical biological types like the system postulated by the process structuralists Goodwin and Webster. Wagner and his collaborators have tried to model selective processes that favor the emergence of discrete developmental "modules" that are stabilized against various perturbations of the developmental system. Although these modules function as developmental invariants in at least some timescales, persisting with apparent disregard for the "conditions of life," they have themselves emerged as a result of the evolutionary process, and they will possess a phylogenetic signature—an association with a particular lineage.

## ACKNOWLEDGMENTS

I have benefited from the comments of Rob Wilson, Gunther Wagner, and Karola Stotz on an earlier draft of this paper.

## NOTE

1. She could exist even if domestic cats had speciated some generations ago—making her, on cladistic principles, a member of one of two new species (LaPorte 1997).

## REFERENCES

Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies* 61, 127–148.

- Brooks, D. R., and D. A. McLennan (1991). *Phylogeny, ecology and behaviour*. Chicago: Chicago University Press.
- Darwin, C. (1859, reprint 1964). *On the origin of species: A facsimile of the first edition*. Cambridge, Mass.: Harvard University Press.
- Fodor, J. A. (1974). Special sciences. *Synthese* 28, 77–115.
- Ghiselin, M. T. (1974a). *The economy of nature and the evolution of sex*. Berkeley: University of California Press.
- Ghiselin, M. T. (1974b). A radical solution to the species problem. *Systematic Zoology* 23, 536–544.
- Gilbert, S. F., J. M. Opitz, and R. Raff (1996). Resynthesising evolutionary and developmental biology. *Developmental Biology* 173, 357–372.
- Goodman, N. (1954). *Fact, fiction and forecast*. London: Athlone Press, University of London.
- Goodwin, B. C. (1994). *How the leopard changed its spots: The evolution of complexity*. New York: Charles Scribner & Sons.
- Goodwin, B. C., and P. Saunders (1989). *Theoretical biology: Epigenetic and evolutionary order from complex systems*. Edinburgh: Edinburgh University Press.
- Goodwin, B. C., S. A. Kauffman, and Murray, J. D. (1993). Is morphogenesis an intrinsically robust process? *Journal of Theoretical Biology* 163, 135–144.
- Gould, S. J., and R. Lewontin (1979). The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London* 205, 581–598.
- Griffiths, P. E. (1992). Adaptive explanation and the concept of a vestige. In P. E. Griffiths, ed., *Trees of life: Essays in philosophy of biology*. Dordrecht, Netherlands: Kluwer.
- Griffiths, P. E. (1994). Cladistic classification and functional explanation. *Philosophy of Science* 61(2), 206–227.
- Griffiths, P. E. (1996a). Darwinism, process structuralism and natural kinds. *Philosophy of Science* 63(3), S1–S9.
- Griffiths, P. E. (1996b). The historical turn in the study of adaptation. *British Journal for the Philosophy of Science* 47(4), 511–532.
- Griffiths, P. E. (1997). *What emotions really are: The problem of psychological categories*. Chicago: University of Chicago Press.
- Griffiths, P. E., and R. D. Gray (1994). Developmental Systems and Evolutionary Explanation. *Journal of Philosophy* 91(6), 277–304.
- Hacking, I. (1991). A tradition of natural kinds. *Philosophical Studies* 61, 109–126.
- Harvey, P. H., and M. D. Pagel (1991). *The comparative method in evolutionary biology*. Oxford and New York: Oxford University Press.
- Hull, D. L. (1976). Are species really individuals? *Systematic Zoology* 25, 174–191.
- Hull, D. L. (1978). A matter of individuality. *Philosophy of Science* 45, 335–360.
- Hull, D. (1986). On human nature. *Proceedings of the Philosophy of Science Association* 1, 3–13.
- Jackson, F. C., and P. Pettit (1988). Functionalism and broad content. *Mind* 97, 381–400.
- Kauffman, S. A. (1993). *The origins of order: Self-organisation and selection in evolution*. New York: Oxford University Press.
- Keil, F. C. (1989). *Concepts, kinds and cognitive development*. Cambridge, Mass.: Bradford Books, MIT Press.

- Kingsland, S. (1985). *Modeling nature*. Chicago: University of Chicago Press.
- Kornet, D. J. (1993). Permanent splits as speciation events: A formal reconstruction of the inter-nodal species concept. *Journal of Theoretical Biology* 164, 407–435.
- Kripke, S. (1980). *Naming and necessity*. Cambridge, Mass.: Harvard University Press.
- LaPorte, J. (1997). Essential membership. *Philosophy of Science* 64(1), 96–112.
- Lewontin, R. C. (1983). The organism as the subject and object of evolution. *Scientia* 118, 65–82.
- Lycan, W. G. (1990). The continuity of levels of nature. In W. G. Lycan, ed., *Mind and cognition: A reader*. Oxford: Blackwell.
- Mayr, E. (1959, reprint 1976). Typological versus populational thinking. In E. Mayr, ed., *Evolution and the diversity of life*. Cambridge, Mass.: Harvard University Press.
- Mayr, E. (1963). *Animal species and evolution*. Cambridge, Mass.: Harvard University Press.
- Mishler, B. D., and R. N. Brandon (1987). Individuality, pluralism and the phylogenetic species concept. *Biology and Philosophy* 2, 397–414.
- Morss, J. (1992). Against ontogeny. In P. E. Griffiths, ed., *Trees of life: Essays in philosophy of biology*. Dordrecht, Netherlands: Kluwer.
- Müller, G. B., and G. P. Wagner (1991). Novelty in evolution: Restructuring the concept. *Annual Review of Ecology and Systematics* 22, 229–256.
- Putnam, H. (1975). The meaning of “meaning.” *Mind, language and reality: Philosophical Papers*, vol. 2. Cambridge: Cambridge University Press.
- Quine, W. V. O. (1977). Natural kinds. In P. Schwartz, ed., *Naming, necessity and natural kinds*. Ithaca, N.Y.: Cornell University Press.
- Riedl, R. (1977). *Order in living systems*. London: Wiley.
- Schank, J. C., and W. C. Wimsatt (1986). Generative entrenchment and evolution. *Philosophy of Science Association* 2, 33–60.
- Smart, J. C. C. (1963). *Philosophy and scientific realism*. London: Routledge & Kegan Paul.
- Sober, E. (1980, reprint 1994). Evolution, population thinking and essentialism. In *From a biological point of view*. Cambridge: Cambridge University Press.
- Wagner, G. P. (1994). Homology and the mechanisms of development. In B. K. Hall, ed., *Homology: The hierarchical basis of comparative biology*. New York: Academic Press.
- Wagner, G. P. (1996). Homologues, natural kinds and the evolution of modularity. *American Zoologist* 36, 36–43.
- Wagner, G. P., G. Booth, and Homayoun, B. C. (1997). A population genetic theory of canalization. *Evolution* 51(2), 329–347.
- Webster, G., and B. C. Goodwin (1996). *Form and transformation: Generative and relational principles in biology*. Cambridge: Cambridge University Press.
- Wimsatt, W. C. (1976a). Complexity and organisation. In M. Grene and E. Mendelsohn, eds., *Boston studies in philosophy of science*, vol. 27: *Topics in philosophy of biology*. Dordrecht, Netherlands: Reidel.
- Wimsatt, W. C. (1976b). Reductionism, levels of organisation and the mind/body problem. In G. Globus, G. Maxwell, and I. Savodnik, eds., *Consciousness and the brain*. New York: Plenum Press.
- Wimsatt, W. C. (1986). Developmental constraints, generative entrenchment and the innate-acquired distinction. In W. Bechtel, ed., *Integrating scientific disciplines*. Dordrecht, Netherlands: Martinus Nijhoff.

# **IV Species in Mind and Culture**

This excerpt is provided, in screen-viewable form, for personal use only by members of MIT CogNet.

Unauthorized use or dissemination of this information is expressly forbidden.

If you have any questions about this material, please contact [cognetadmin@cognet.mit.edu](mailto:cognetadmin@cognet.mit.edu)