Natural Kinds and Induction in the Special Sciences*

[Author Suppressed for Blind Review]

March 10, 2009

Abstract

This paper discusses a strategy of arguing for causal theories of natural kinds based on considerations of reliable induction, as pursued by Boyd, Kornblith, and Millikan. It argues that, even if the strategy is successful for natural kinds that appear in exceptionless laws, it fails for the laws articulated by the special sciences because these only hold with exceptions, or *ceteris paribus*.

Keywords INDUCTION · NATURAL KINDS · SPECIAL SCIENCES · HOMEOSTATIC PROPERTY CLUSTERS · CETERIS PARIBUS LAW

1 Introduction

Alongside straightforward empirical discovery, the sciences also exhibit significant conceptual innovation. One way of classifying phenomena is replaced by another, perhaps piecemeal, perhaps wholesale. That such replacement takes place suggests that one classificatory scheme can be better suited to the needs of a discipline than another, and a theory of natural kinds seeks to tell us what makes that so. In this paper, I want to discuss theories that tie the suitability of natural kinds very closely to inductive inference: natural kinds must support them. This basic thought goes back at least as far as Quine (1969), but I am particularly interested in theories that accord causal notions a central role in constraining natural kinds, such as those advanced by Boyd, Kornblith, and Millikan.[1] The larger question I want to investigate concerns a contrast

*[Acknowledgments suppressed for blind review.]

[1] See, e.g., Boyd (1988, 1991, 1999a,b), Kornblith (1993), and Millikan (1999a).

between what might be termed atomic as opposed to systemic approaches to natural kinds. On an atomic approach, we derive constraints on natural kinds from features of generalizations considered individually, while on a systemic approach, we derive these constraints from features of generalizations considered collectively. My arguments in this paper are aimed at establishing that a plausible atomic approach to natural kinds in the special sciences fails and I hope to thereby support a more systemic outlook.

Specifically, I will argue that the strategy of deriving substantive constraints on natural kinds by connecting them to reliable induction fails as applied to the special, i.e., non-fundamental sciences, such as biology, psychology, economics, geology, or meteorology. The problem I will raise turns crucially on the fact that these sciences articulate non-strict generalizations, generalizations that are usually called *ceteris paribus* laws. Because of this fact, the only constraints on special science kinds we can argue for on the basis of the connection between natural kinds and induction are trivial.[2]

§2 introduces the theories I discuss and shows the argumentative strategy at work in laws that take the form of universal generalizations. §3 contains my main argument: I begin by arguing that the most direct way of extending the strategy to non-strict laws fails, and I also consider and reject some more sophisticated extensions. §4 discusses some of the extensions and implications of those arguments.

2 Induction, Natural Kinds, and Causal Homogeneity

Suppose it true that all *A*s are *B*. Under what conditions is it appropriate to infer that this is so on the basis of the knowledge that a non-exhaustive sample of *A*s consists only of *B*s? This is one of the question we can ask when we wonder whether the predicate *B* is projectable with respect to the class of *A*s. Famously, 'is a silver coin' is not projectable with respect to the class of coins in my pocket. Even if all of these coins are silver, I may not conclude that this is so on the basis of anything short of inspecting all of them.[3] The theories of natural kinds that I discuss offer a theory of

[2]I want to emphasize that my arguments do not turn on the concern that cp-laws are somehow defective, a concern that is usually pressed by asking whether they themselves might be trivially true. For relevant literature, see Earman and Roberts (1999); Earman et al. (2002); Fodor (1991); Glymour (2002); Lange (1995, 2002); Mott (1992); Schiffer (1991); Schurz (2001, 2002); Silverberg (1996); Woodward (2002). I've argued in detail that this worry is misplaced. See my[Reference suppressed for blind review].In any case, I am concerned only to argue that the constraints we can derive are trivial, not that the laws from which they might be derived are trivial.

[3]I say that a predicate is projectable "with respect to a class" because obviously, there are some (natural) classes with respect to which 'is a silver coin' is projectable. Thus, more generally we should always ask whether a predicate is projectable with respect to this or that class, not whether it is projectable *simpliciter*.

projectability and seek to thereby argue for a further constraint on natural kinds. Here are some representative statements of the approach, beginning with Hilary Kornblith.

> How [. . .] does the existence of natural kinds help to explain what it is about the world that makes inductive knowledge of it possible? The existence of natural kinds brings with it a certain causal structure. Only certain combinations of properties may be found together in a single individual, for not all combinations are causally possible or causally stable. [. . .] I argue that natural kinds make inductive knowledge of the world possible because the clustering of properties characteristic of natural kinds makes inferences from the presence of some of these properties to the presence of others reliable.[4]

More starkly, Boyd says that

> [i]t is a truism that the philosophical theory of *natural* kinds is about how classificatory schemes come to contribute to the epistemic reliability of inductive and explanatory practices.[5]

He goes on to give an example of natural kinds at work.

> Suppose that you have been conducting experiments in which you exposed various salts of sodium to flames. In each of many cases, the flame turned yellow. You conclude that always (or almost always) if a salt of sodium is heated in a flame, then a yellow flame results. You are right, and your inference is scientifically respectable.
>
> Your inductive success in this matter is a reflection of the fact that the categories *salt of sodium*, *flame*, and *yellow* are natural categories in chemistry, and of the fact that the hypothesis you formulated with the aid of reference to these categories is a projectable one.
>
> [. . .] Your deployment of projectable categories and generalizations allowed you to identify a *causally sustained* generalization.[6]

While Kornblith and Boyd hold a monolithic view according to which a single theory of natural kinds applies to all sciences, Millikan gives divergent theories of projectability

[4] Kornblith (1993, 6-7).
[5] Boyd (1999a, 146), emphasis in the original.
[6] Boyd (1999a, 147), emphasis in the original.

for fundamental sciences on the one hand and special sciences on the other. Thus, she says quite generally that

> in the case of a proper law, there is a reason why the corresponding empirical generalization holds, which reason lies in the natures of the antecedent and consequent of the law, rather than the accidental positioning of exemplifications of these conditions along with other things in the historical order. Similarly, [. . .] being projectable is a matter of the intrinsic nature of a property, to be distinguished from accidental continued coincidences in its occurrences.[7]

In the case of the special sciences, what allows a kind to figure in such a non-accidental generalization is the fact that it is what she calls a "historical kind:"

> The members of these kinds are like one another because of certain historical relations they bear to one another. [. . .] [E]ach exhibits the properties of the kind because other members of that same historical kind exhibit them. Inductions made from one member of the kind to another are grounded because there is a certain historical link between the members of the kind that causes the members to be like one another.[8]

Let me extract what is common to all of these positions. Their authors all pursue a two-part strategy, arguing for constraints on natural kinds from constraints on inductively supportable laws about these kinds. They all agree that an inductively supportable law needs to be causally supported. That is to say, in order for the inductive inference from an observed sample of *A*s, all of which are *B* to the conclusion that *all As are B* to be appropriate, the two properties need to be causally related. They may be related in virtue of the fact that, in any case in which the two properties are coinstantiated, the presence of the one causes the presence of the other, or the presence of both is the effect of a common cause; there may be other causal structures, as well.[9] Crucially,

[7]Millikan (1999a, 47-8).

[8]Millikan (1999a, 54-5). One might think that appeal to historical relations, which are extrinsic, does not fit into the constraints imposed by the earlier quote, since there Millikan says that being projectable is a matter of the *intrinsic* nature of a property. However, as Millikan concedes in her response to Boyd's comments on this passage, she meant "intrinsic" to simply mean "non-accidental." See Millikan (1999b).

[9]This schematic representation is too narrow, since there are generalizations that relate propery instantiations without requiring the properties to be coinstantiated. An example concerns the effect of one massive body on another, that every massive body is such that other massive bodies in the vicinity experience a gravitational force of such-and-such magnitude. So to be more precise, the main text should read that in all cases in which the properties picked out by *A* and *B* are related in the way the generalization says they are, there is a single mechanism that accounts for their being so related. However, I remain with the more narrow formulation to ease exposition. Nothing else turns on the difference.

however, it needs to be the case that whatever the relevant causal structure is, it is the same in all cases of coinstantiation. This last requirement is what distinguishes the merely accidental from the causally supported generalizations.

The authors I've cited differ on why the generalizations need to be causally supported in this sense. Millikan seems to argue that projectability requires non-accidentality, and offers her theory of causal support as an analysis of non-accidentality. Boyd and Kornblith argue for the causal support requirement by arguing that inductive inferences are reliable only if they issue in such causally supported conclusions. But that difference need not concern us here—all that matters for my purposes is that the generalizations be causally supported in the sense I described.

One may, of course, be worried about this part of the argument. Perhaps the reliability condition on appropriate inductive inference is mistaken; perhaps the reliability condition is itself correct, but need not be spelled out in the causal terms favored by Boyd and Kornblith; or perhaps Millikan's analysis of non-accidentality in broadly causal terms is mistaken.[10] However, I want to simply grant that the argument is sound up to this intermediate conclusion and focus on the next step.

That next step is to argue that since inductively supportable generalizations must be causally supported, we can draw a further conclusion about the properties mentioned in the generalization. Causal support requires that in each case in which the property of being an *A* and being *B* is coinstantiated, there is a single mechanism that accounts for this. Moreover, since all *A*s are *B*, it follows that all *A*s are involved in this mechanism. That is to say that the *A*s themselves are what we may call *causally homogeneous* with respect to being *B*: all *A*s are involved in the same process when it comes to determining whether they are *B*.

How should we relate the notion of causal homogeneity to a general constraint on natural kinds? Depending on our choices, we will arrive at a stronger or weaker theory of natural kinds. There are two parameters to consider. The first is whether a potential natural kind has to be causally homogeneous with respect to all of the properties that appear in true generalizations about the kind or not. Let me illustrate the issue, remaining with the example Boyd introduced regarding salt of sulphor. Such salt is mentioned in several true generalizations, not just about the color of the flame it produces when burned, but also about characteristic smells or what it reacts with. In order for the property of being salt of sodium to be a natural kind, we can require that it is causally homogeneous with respect to all of these properties, with respect to at least

[10]See, e.g., Vogel (1987, 2007) and citations therein.

one, or something in between. The second parameter concerns the scope of the theory of natural kinds itself. Is the constraint on natural kinds we are motivating by considering the requirements of appropriate inductive inference one that *all* natural kinds need to satisfy, only some of them, or something in between?

Because of the basic idea behind the inductive reliability approach, the most interesting version of the position I am discussing results from making the claims maximally strong. Most plausibly, that approach should extend to every generalization that can be inductively supported, which is just to say, every generalization. That means that every true generalization about a kind should be relevant for determining whether that kind is natural, and we should require the theory to be about all kinds. Proponents of the inductive reliability approach should therefore endorse constraints that instantiate the following schema. For all properties *K*, *K* is a natural kind only if a suitable subclass of *K*s is causally homogeneous with respect to all of the properties that appear in true generalizations about *K*s. Specific constraints are derived once we've spelled out what the suitable subclass of *K*s is. If we focus on potential natural kinds that appear only in true, universal generalizations, then the suitable homogeneity constraint we can impose on these kinds is very strong, indeed.

[STRONG HOMOGENEITY] For all properties *K*, *K* is a natural kind only if for each property *P* that appears in a true generalization about *K*s, all *K*s are causally homogeneous with respect to *P*.

It is STRONG HOMOGENEITY that is most plausibly at work when we consider the famous jadeite/nephrite example. The material we ordinarily call 'jade' is made up of two different materials, jadeite and nephrite, and because the property of being jade is not causally homogeneous with respect to any of the properties investigated by chemistry, it is not a natural kind for chemistry.

Let me contrast STRONG HOMOGENEITY with a much weaker constraint. First, instead of requiring that a property *K* is a natural kind only if *all* of the *K*s are causally homogeneous with respect to the properties that appear in true generalizations about *K*s, we only require that *some* of them are. Second, once we weaken the requirement this way, a scope issue becomes important that didn't arise so long as all *K*s had to be causally homogeneous. Should we require that there is some subset of *K*s that are causally homogeneous with respect to all of the properties mentioned in true generalizations (i.e., it's the same *K*s for all of the properties) or should we require only that for each of the properties mentioned in a true generalization, there is a subset of *K*s that is homogeneous with respect to that property (i.e., different homogeneous subsets for different properties)? The latter requirement yields the weakest constraint possible, indeed, no constraint at all.

[TRIVIAL HOMOGENEITY] For all properties *K*, *K* is a natural kind only if for each property *P* that appears in a true generalization about *K*s, there is a subset of *K*s that are causally homogeneous with respect to *P*.

I call this TRIVIAL HOMOGENEITY because every property *K* satisfies it. To see the structure of the problem, notice for example that the property of being jade satisfies this constraint because it has causally homogeneous subsets, to wit, one composed of the samples of jadeite and one composed of the samples of nephrite. The burden of this paper is to argue that we can impose at most TRIVIAL HOMOGENEITY on the kinds in the special sciences if we try to adapt the two-stage argumentative strategy I presented above to those disciplines. Discharging this burden is the task of the next section.

3 Non-Strict Generalizations

The initial observation is that all of the special sciences articulate generalizations that we can intuitively characterize as non-strict. Slightly more precisely, the special sciences articulate generalizations that can be turned into falsehoods by introducing quantifiers with universal force into their formulation. 'Ravens are black,' 'humans learn to speak by age five,' and 'snow is white' are all true even though some ravens aren't black, some humans don't learn to speak by age five, and some snow isn't white. More generally, the special sciences articulate generalizations of the form *As are B*, which can be true even when *all As are B* is false.[11] We can already see why we cannot impose STRONG HOMOGENEITY as a general constraint on natural kinds. Focus on 'ravens are black.' If all ravens were causally homogeneous with respect to being black, they would all be black. Since they aren't, they aren't.

[11] I do not use the expression 'ceteris paribus' to pick out the non-strict generalizations of the special sciences because that locution is not systematically used outside of economics—see Schiffer (1991); Woodward (2002). Using sentences with bare plurals, i.e., plural subjects without a determiner like 'all,' is much more natural. These generalizations still exhibit the important features usually ascribed to cp-laws, since they tolerate exceptions and have modal import.

3.1 Ascribing Dispositions

Before going any further, let me set aside one class of generalizations that are clearly non-strict in the sense I just introduced, but that aren't troublesome for the proponent of the inductive reliability strategy. In fact, the generalization Boyd mentioned in his quote is one of them. He says that on the basis of observing yellow-burning salts of sodium, we may conclude that "always (or almost always) if a salt of sodium is heated in a flame, then a yellow flame results." The parenthetical qualification isn't idle, since there are obviously situations in which one has a bit of salt of sodium, heats it in a flame, and yet the flame does not turn yellow. The salt of sodium might have been mixed with some other chemical, for instance, so that the flame of the compound is green. In order to have a generalization that isn't refuted by the possibility of these kinds of cases, the qualification is needed.

This is an instance of a very general pattern involving ascriptions of dispositional properties. When we say that salt of sodium burns yellow, we are ascribing a disposition to that salt. And like all dispositions, its manifestations can be blocked or masked by various aspects of a situation that would otherwise be one in which the disposition was triggered.[12] That the manifestation is so masked does not mean that the object or sample does not have the disposition ascribed. Compare: a carefully packaged glass is fragile, even if it happens not to break when it is dropped because of that packaging.

The crucial feature of these non-strict generalizations is that they are non-strict insofar as we consider the possible triggers and manifestations of the disposition ascribed. However, they may well be utterly strict with respect to the things to which the disposition is ascribed. That is to say, the generalization 'salt of sodium burns yellow when heated in a flame' is one that covers all salt of sodium, though it does not cover every single instance in which a sample of such salt is heated in a flame. That means that the initial argument I reconstructed on behalf of the inductive reliability approach goes through without a hitch so long as we confine ourselves to universal generalizations about dispositional properties. As far as these cases are concerned, STRONG HOMOGENEITY may be correct.

However, not all non-strict generalizations are ascriptions of dispositions. 'Ravens are black,' 'grass is green,' and 'snow is white' are all non-strict, but they are not ascriptions of dispositions. It is simply false, for example, that albinos are disposed to be

[12]There is a large literature that discusses such masking and the problems it poses for counterfactual analyses of dispositions. See Lewis (1999) for a popular statement that many theorists respond to, and Fara (2005) for an overview of responses to the problem, along with his own contribution.

black but that, for some reason, their disposition is kept from being manifested. Hence, we cannot assimilate all non-strict generalizations to strict ascriptions of dispositions.

Since without this claim the remainder of my arguments is moot, let me consider an objection at this point. One might disagree with me, holding that 'ravens are black' does ascribe a disposition, after all, arguing thus. This claim seems to be properly paraphrased by saying that ravens tend to be black, and claims about tendencies are plausibly paraphrased in terms of dispositions. After all, the uncontroversial ascription of the dispositional property of fragility to a glass is well glossed by saying that this glass tends to break when dropped.

The parallel between the blackness of ravens and the fragility of glasses is only superficial, however, since in the one case, we are making a judgment about the probability of finding a black raven in the population as a whole, while in the other, we're making a claim about a disposition. An argument: if we're really ascribing a disposition, then the same property we're ascribing in the course of formulating a generalization is one we should just as happily ascribe to an individual. However, if we're ascribing a property that only populations can have, then the ascription of that same property to an individual should be awkward. With this in mind, the obvious difference between 'Al (the albino raven) tends to be black' and 'this glass tends to break when dropped' shows that 'ravens are black' is not an ascription of a disposition.

Finally, the mere fact that a generalization ascribes a disposition does not, in itself, ensure that it is universal with respect to the class of objects to which the disposition is ascribed. Consider, for example, 'humans learn to speak by age five.' This generalization ascribes the disposition to learn a human language by a certain age, and like all dispositions, its manifestation can be masked or otherwise derailed. But there are also some children born without this disposition. So generalizations can be non-strict twice over, and to the extent that they are non-strict with respect to the objects generalized over, they pose a problem for the inductive reliability approach.

So while it's true that some non-strict generalizations need not worry the proponent of the inductive reliability approach to natural kinds because these are ascriptions of dispositions, others cannot be so easily handled. It is these generalizations I want to discuss in greater detail now.

3.2 Majorities

The crucial question for my opponents is whether they can impose a stronger constraint than TRIVIAL HOMOGENEITY, given that they definitely cannot impose STRONG HO-

MOGENEITY. One might think that they can. After all, many non-strict generalizations about a kind aren't just about *some* members of that kind. They seem to make a significantly stronger claim, while still falling short of a fully universal one. This is already implicit in Boyd's gloss on the ascription of a disposition in terms of "almost always," which suggests that though the disposition isn't manifested in every case in which the triggering conditions obtain, it is manifested in almost all of them. And this gloss seems appropriate for other generalizations, as well. 'Ravens are black,' 'humans learn to speak by age five,' and 'snow is white' are all well-paraphrased as claims about the majority of the members of the kind involved. If non-strict generalizations could always be paraphrased in these terms, we could impose a stricter requirement on a potential natural kind property. Not only would there need to be a causally homogeneous subclass among the things that have the property, that subclass would also need to contain the majority of those things.

This proposal has the added benefit of being in the spirit of the initial inductive reliability strategy. Suppose it true that if *A*s are *B*, then most *A*s are *B*. In that case, we can also see why induction is often reliable. Given that most *A*s are *B*, it follows that a randomly chosen *A* is more likely than not to be a *B*, so that if one simply projects the property of such a randomly chosen *A* to the kind in the form of a non-strict generalization, one is more likely than not to arrive at the correct generalization. On this proposal, therefore, the constraints on natural kinds aren't just motivated by a causal analysis of non-accidentality, but are such that, when they are satisfied, induction over these kinds is reliable—not as reliable as in the sciences that arrive at universal generalizations, but reliable for all that.

As before, an appeal to majorities can take a stronger or weaker form. The stronger version results from requiring that there is a single subclass that incorporates the majority of members of the kind and which is causally homogeneous with respect to all of the properties that appear in true generalizations about members of the kind, the weaker from requiring merely that for each of these properties, there is such a subclass. If the stronger requirement is satisfied, so is the weaker. I'll argue now that the weaker requirement is too strong to impose generally, and this argument will serve to argue against the stronger requirement, as well.

The question then is whether the schematic conditional connecting non-strict generalizations to claims about the majority of members of a kind is true. In evaluating this question, we need to be a little bit careful. Taken as a claim about what is true at any moment in time, it is clearly false. Suppose that by some temporary fluke, most ravens happen to be white. Almost all ravens that ordinarily would have been black were painted.[13] Such momentary fluctuations would not incline us to reject 'ravens are black,' though it would clearly be false that most ravens are black. To deal with these kinds of flukes, we should interpret the connection between a non-strict generalization about some kind and a generalization about the majority of members of that kind as concerning the long run. For example, if ravens are black, then over the suitably long run, most ravens are, too.

However, even taking long run patterns into account, it is not in general true that if *A*s are *B*, then most *A*s are *B*. There are many cases in which we want to accept a non-strict generalization even though the cases that conform to it are consistently in the minority. Non-strict generalizations about reproduction serve to make the point. It's true that ravens lay eggs and that humans give birth to live young, but in each case, it is a completely stable fact that only a minority of the members of the kind conform to the generalization: neither males nor immature or infertile females do. Likewise, generalizations about sexual selection pose problems. It is true that lions have manes, that cardinals are red, and that birds of paradise have astounding plumage. But in each case, only the mature males exhibit these features, and thus it is false that most members of the kind conform to the generalization, even over the long run.

3.3 Excluding Irrelevant Exceptions

One might respond to my objection that in the cases I consider, a mere minority of members of the kind conform to the generalization because I am including in the scope of that generalization members that are intuitively irrelevant to it. Of course most humans don't give birth to live young, but who would have thought that males extrude offspring in any way at all in the first place? For all that I've said, the rebuttal goes, the conditional that if *A*s are *B*, then most *A*s are, still holds. We simply have to take care to exclude irrelevant *A*s from consideration when we evaluate the consequent. I want to take the sense underlying this rebuttal seriously by offering a general diagnosis of what makes these particular counter-examples work, along with what seems to me to be the best way of systematically avoiding them. But as I'll also argue, there are other problems for the 'most'-conditional that cannot be handled by these means. That conditional, even in the altered form I'll present here, cannot be maintained.

[13]One might protest at this stage that *in the relevant sense*, ordinary ravens that have been painted white still count as black, taking inspiration from Austin (1975), and following him Travis (1985, 2000). However, the same point could be made by considering number of legs: 'ravens have two legs' is true even if most ravens happen to have lost a leg in an accident. The point still holds.

The core idea that animates my proposal is this.[14] A non-strict generalization such as ‘ravens lay eggs’ is compatible with the existence of ravens that do not lay eggs. Let’s call all of these ravens exceptions to the generalization. Among the exceptions, we can draw a further distinction. There are some that were never really in the running in the first place while the rest were. No matter how things turn out to be with members of the former class, it would not affect our evaluation of the generalization one bit, and that makes these exceptions strictly speaking irrelevant. In my example, the male ravens fall into that category. Non-strict generalizations implicitly exclude such irrelevant counterexamples from their purview, and for that reason, we need to exclude them as well when we evaluate whether most members of the kind conform to the generalization. That is, the thesis my opponents should defend is this: if *A*s are *B*, then *among the relevant A*s, most *A*s are *B*. Call this the *relevance conditional*.

To see how this works, let me consider a couple of examples. First, the relevance conditional does well for ‘ravens lay eggs’ because only females are relevant to its evaluation, and among the females, the majority does indeed lay eggs. We can also illustrate the need to consider both the relevant cases, and even among the relevant ones, to offer a generalization that is weaker than a fully universal one. Sharks generally reproduce by producing live young (though not with a placenta, but in an egg that hatches inside the female shark, a process called ovoviviparity. However, some sharks have reproduced by parthenogenesis (asexually). Informally, if we want to evaluate ‘sharks reproduce by ovoviviparity,’ we exclude as irrelevant the male sharks, leaving us with the females. But that includes sharks that have reproduced by parthenogenesis, and hence it is true that most, though not all, of the relevant sharks reproduce by ovoviviparity.

In the case of the simpler conditional connecting non-strict generalizations to generalizations about the majority of the members of the kind, we could formulate stronger or weaker constraints based on that conditional. In the case of the relevance conditional, by contrast, we can only formulate a weaker constraint, since it is built into the conditional that which *A*s are in the scope of the non-strict generalization depends on which generalization we’re interested in. To take an extreme example, ‘lions have manes’ concerns only male lions, ‘lions suckle their young’ only female ones. Hence, it’s ruled out from the start that the same *K*s could be causally homogeneous with respect to all of the properties mentioned in true generalizations about *K*s.

[14]The outlines of the proposal I discuss are drawn from the work of Ariel Cohen on non-strict generalizations in natural language, sentences that are usually called ‘generics.’ See Cohen (1999a,b, 2004).

Nonetheless, we can still formulate an interesting constraint based on the relevance conditional. In this case, the constraint has to be placed on whole collections of kinds at a time, rather than on kinds considered one-by-one. Roughly, such a constraint might say that a property is a natural kind only if there is a subset of the *K*s, *K**, such that *K** is causally homogeneous with respect to some property *P* that appears in a true generalization about *K*s, and there is a subset of the *K*s, *K*† that results by focusing on the *K*s that also have some other natural kind properties, and *K** comprises the majority of members of *K*† that are also *P*. Informally, what's required is that there be a subset that is causally homogeneous with respect to a property that appears in a true generalization, and that makes up the majority of *K*s that have one or more further natural kind properties in common.[15]

One might have two worries about this development of the intuition that when we're evaluating the connection between a non-strict generalization about *K*s and what's true of the majority of *K*s, we should only focus on the relevant *K*s. Both concern the restriction to relevant members of the kind in the formulation of the relevance conditional. First, one needs to ensure that the non-strict generalizations themselves don't turn into trivial truths, claims of the form *A*s that are *F* are *F*, for example, 'ravens that are egg-layers are egg-layers.' A full development of the strategy I'm offering on behalf of the inductive reliability theorist would presumably include restrictions on how the relevant members of the kind are determined. I'll simply assume that this can be done.

More importantly, one might also worry that the restriction to the relevant members makes not the non-strict generalization itself, but the relevance conditional, trivial. Specifically, one might worry that, with an artfully chosen circumscription of the relevant *A*s, it's always true that if *A*s are *F*, then most of the relevant *A*s are *F*. This concern isn't just important for a potential development of a view incorporating the relevance conditional, it also makes the dialectical situation somewhat touchy. Since I want to argue that there are counter-instances to the relevance conditional, I need to make it plausible that in the cases I'll citen, my opponent cannot simply counter with a particularly favorable way of restricting herself to the relevant cases. I'll deal with that in two ways. One kind of example I'll mention depend on the intuitive sense that there's really only one natural way of specifying the relevant cases, and on that specification, the example is troublesome for my opponent. The other kind of example

[15] Such structural constraints may or may not be trivial. It is perhaps always possible to begin with an an arbitrary property *K* and then cook up other properties that can serve to show that *K* stands in the relevant structural relations. For the purposes of this paper, I will simply waive this concern, since I want to argue that the relevance conditional is false in any event.

depends not on what the relevant cases are, but only on relationships among relevant generalizations.

Here, then, are the counterexamples. There are cases in which, even once we have narrowed our focus to the most reasonable relevant subclass, it is still the case that only a minority of the members of that class conforms to the generalization at issue. For example, it is true that sea-turtles are long-lived, even though most turtles die young, indeed, almost immediately after they hatch. In this case, the most natural way of picking out the relevant class of sea-turtles are the ones that have some life-span or other, and that just includes all of the sea-turtles. Obviously, we might want to say that in this case, we're really only interested in those sea-turtles that reach an age that reflects their potential longevity, but that way of restricting ourselves to relevant cases threatens to trivialize the claim.

As I said, relying on examples like this is dialectically touchy because there is some room for alternative ways of specifying the relevant class of objects for the purposes of evaluating the relevance conditional. So consider instead sets of generalizations like the following. 'Elephants live in Africa and elephants live in Asia,' or 'bears live in North America, South America, Europe, and Asia,' and focus specifically on the first of these. Suppose that whatever makes an elephant relevant for evaluating the first conjunct 'elephants live in Africa' is some restriction R. The crucial claim for my argument is that whatever R is, it should be the very same restriction that's at issue in the second conjunct. And if that's true, then one of the conjuncts has to be a counterexample to the relevance conditional. To see this, imagine that the first conjunct conforms to the relevance conditional. In that case, most elephants that satisfy R live in Africa. But it then follows that it's false that most elephants that satisfy R live in Asia, thus making the second conjunct a counterexample. The situation is symmetrical, so we definitely have a counterexample either way. Other counterexamples that have the same structure concern the color of cats (white, black, orange, tabby, blue, gray, etc.) the color of roses (red, white, yellow), and so on. And in these cases, it doesn't matter what exactly the restriction is. All that matters is that for each of the sentences in such a conjunction, the restriction is the same.

Finally, it will not help my opponent to protest that there are other ways of describing the same situation. She might say, for example, that the habitats of elephants and bears are better described in terms of modal locutions: 'elephants can live in Africa and Asia,' or 'roses can be red, white, and yellow.' The point of such a reformulation is that where before, we were ascribing incompatible properties to various groups of elephants (living in Africa and living in Asia), one of which had to be a minority, we are now ascribing a single modal property to most elephants, *simpliciter*, the modal property of possibly coming from Africa and or Asia. But it is not true that those members of the kind that aren't actually from, say, Africa could just as well come from there, anyway. Asian elephants couldn't, in the relevant sense of 'could', come from Africa. The point is perhaps more obvious for bears: polar bears couldn't come from the Asian jungles. We could catch some and put them there, but the relevant sense of 'can' in these example involves what is compatible with the natural habitat of bears, and the jungles of Asia aren't a natural habitat for polar bears. Thus, the relevance conditional fails, as well.

I take it that an appeal to the relevance conditional, and the structural constraint on natural kinds that comes in its wake, is the best strategy for proponents of the inductive reliability approach to natural kinds. Thus, the failure of the relevance conditional brings with it the failure of that approach.

4 Upshots

The arguments I have made are aimed at establishing that we cannot, in general, derive non-trivial causal homogeneity constraints on natural kinds in the special sciences by appealing to considerations of what is required for reliable induction. The problem is that the only constraint on kinds to which there aren't any counterexamples is TRIVIAL HOMOGENEITY: *K* is a natural kind only if , for each property *P* that appears in a true generalization about *K*s, there is a subset of *K*s that is causally homogeneous with respect *P*. Attempts to impose a stronger, fully general condition are frustrated by the existence of some extremely weak generalizations.

If these arguments succeed in showing that considerations of inductive reliability only motivate a trivial constraint on natural kinds in the special sciences, then they can be extended to show that considerations of explanatory power can do no better. To make that argument, I need to say something about the role of generalizations in explanation. Given that my opponents want to establish that natural kinds need to satisfy broadly causal conditions, they should endorse a broadly causal account of explanation. On such an account, information about some explanandum is explanatory only if it is causal information. In the case of explaining why an event occurs, for example, explanatory information needs to yield information about the causal history of that event. One might then hold that generalizations enter causal explanation to the extent that they capture important patterns in the causal histories of the events mentioned in the explanation.[16] Why are polar bears so well suited to their environment? A good explanation could mention (*inter alia*) that polar bears are white and point out that white bears are well camouflaged in the normal habitat for polar bears.

These considerations suggest that if we want to draw conclusions about natural kinds by thinking about the explanatorily useful generalizations they can appear in, we can only, in the first instance, impose a condition on these generalizations. As before, we can then try to derive constraints on the natural kinds that can appear in such explanatorily useful generalizations, but any such argument will face exactly the same limitations as the arguments that motivate causal constraints on the relevant generalizations via consideration of inductive reliability. In both cases, the extent to which we can derive a constraint on natural kinds the structure of the argument will be exactly as wide as the scope of these generalizations. In neither case can we impose a constraint on members of a kind that aren't in the scope of a generalization based on consideration of that generalization.

At this point, the proponent of basically causal theories of natural kinds has two responses available. The first takes issue with my contention that an approach to natural kinds that takes off from considerations of inductive reliability must treat all generalizations equally. Someone who pursues this avenue will want to say that the notion of natural kind itself is a concept with an open texture, so that no single theory of natural kinds can hope to apply to all. Hacking (1991) has offered this view as applied to categories in general. However, the response to my arguments we are considering right now doesn't just hold that what makes the categories of cookbooks and cricket good is different from what makes the categories of chemistry good. This response is committed to the more remarkable thesis that even within an arena of inquiry, such as a single scientific discipline, different theories of natural kinds are appropriate to different particular kinds.

One reason to be suspicious of this pluralism about naturalness is that it is a contingent fact about the world that certain natural kinds are such that, even among the relevant members of the kind, only a small minority conform to the generalization, and hence only a small minority have to be causally homogeneous with respect to the property predicated. Think about the example of the sea-turtles again in this light. Sea-turtles participate in a stable predator-prey dynamic, but that fact itself is highly contingent. The predators might go extinct and it might be the case that most sea-turtles

[16]See, for example, Lewis (1986); Salmon (1989); Strevens (2004).

now die at an age that more closely reflects their constitution. It seems highly implausible that along with the change in predation patterns, we would also see a change in what makes the property of being a sea-turtle a natural kind. That in turn suggests that whether or not most members of a kind conform to a non-strict generalization should not make a difference to what makes the kind mentioned in the generalization natural in the first place.

An alternative response to my arguments is to think that examples like the ones I've cited in this paper are true generalizations only because they fit into a larger system of generalizations about the kind at issue. On this approach, the reason that 'sea-turtles are long-lived' is true is not that the causal mechanism that causally supports the generalization has important theoretical features considered in isolation—that, for example, an inductive inference to this conclusion is particularly reliable. Instead, the reason the generalization is true is that the causal mechanism bears theoretically important relations to mechanims that underwrite other generalizations about sea-turtles. This kind of account would certainly comport well with the suggestion that whether most sea-turtles die young or not is simply irrelevant to what makes the generalization true. It's the underlying causal structure that the generalization gets at that is responsible.

The central feature of such an approach that distinguishes it from the inductive reliability strategy I've considered here is that it considers many generalizations collectively and hence views the utility of kinds in the first place through the prism of how well they support systematic representations of the world, rather than a distinctive way of arriving at the generalizations—induction—or a distinctive use of them—explanation. On this approach, what makes a property a natural kind is the fact that it occupies a causal nexus in a complex web of causal connections.

References

Austin, J. L. *How To Do Things With Words* (Cambridge, MA: Harvard UP, 1975).

Boyd, R. N. "How To Be A Moral Realist". In G. Sayre-McCord, ed., *Essays on Moral Realism*, 181–228 (Ithaca: Cornell UP, 1988).

—. "Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds". *Philosophical Studies*, 61(1-2), (1991), 127–148.

—. "Homeostasis, Species, and Higher Taxa". In R. A. Wilson, ed., *Species: New Interdisciplinary Essays*, 141–185 (Cambridge, MA: MIT Press, 1999a).

—. "Kinds, Complexity and Multiple Realization: Comments on Millikan's *Historical Kinds and the Special Sciences*". *Philosophical Studies*, 95(1-2), (1999b), 67–98.

Cohen, A. "Generics, Frequency Adverbs, and Probability". *Linguistics and Philosophy*, 22, (1999a), 221–253.

—. *Think Generic!* (Stanford, CA: CSLI Publications, 1999b).

—. "Generics and Mental Representation". *Linguistics and Philosophy*, 27(5), (2004), 529–556.

Earman, J. and J. Roberts. "*Ceteris Paribus*, There is no Problem of Provisos". *Synthese*, 118(3), (1999), 439–478.

Earman, J., J. Roberts, and S. Smith. "*Ceteris paribus* Lost". *Erkenntnis*, 57(3), (2002), 281–301.

Fara, M. "Dispositions and Habituals". *Noûs*, 39(1), (2005), 43–82.

Fodor, J. A. "You can fool some of the People All of the Time, Everything Else Being Equal; Hedged Laws and Psychological Explanations". *Mind*, 100(1), (1991), 19–34.

Glymour, C. "A Semantics and Methodology for Ceteris Paribus Hypotheses". *Erkenntnis*, 57(3), (2002), 395–405.

Hacking, I. "A Tradition of Natural Kinds". *Philosophical Studies*, 61(1-2), (1991), 109–126.

Kornblith, H. *Inductive Inference and Its Natural Ground* (Cambridge, MA: MIT Press, 1993).

Lange, M. "Are there Natural Laws Concerning Particular Biological Species". *The Journal of Philosophy*, 92(8), (1995), 430–451.

—. "Who's Afraid of *Ceteris-Paribus* Laws? Or: How I Learned to stop Worrying and Love Them". *Erkenntnis*, 57(3), (2002), 407–423.

Lewis, D. K. “Causal Explanation”. In *Philosophical Papers, Vol. 2*, 214–240 (New York: Oxford UP, 1986).

—. “Finkish Dispositions”. In *Papers in Metaphysics and Epistemology*, 133–151 (Cambridge, UK: Cambridge UP, 1999).

Millikan, R. G. “Historical Kinds and the ”Special Sciences””. *Philosophical Studies*, 95, (1999a), 45–65.

—. “Response to Boyd’s Commentary”. *Philosophical Studies*, 95, (1999b), 99–102.

Mott, P. “Fodor and Ceteris Paribus Laws”. *Mind*, 101(402), (1992), 335–346.

Quine, W. “Natural Kinds”. In *Ontological Relativity and Other Essays*, 114–138 (New York: Columbia UP, 1969).

Salmon, W. C. “Four Decades of Scientific Explanation”. In P. Kitcher and W. C. Salmon, eds., *Minnesota Studies in the Philosophy of Science Vol. 13: Scientific Explanation*, 3–219 (Minneapolis: University of Minnesota Press, 1989).

Schiffer, S. “Ceteris Paribus Laws”. *Mind*, 100(1), (1991), 1–17.

Schurz, G. “Pietroski and Rey on Ceteris Paribus Laws”. *British Journal for the Philosophy of Science*, 52(2), (2001), 359–370.

—. “*Ceteris paribus* Laws: Classification and Deconstruction”. *Erkenntnis*, 57(3), (2002), 351–372.

Silverberg, A. “Psychological Laws and Non-Monotonic Logic”. *Erkenntnis*, 44(2), (1996), 199–224.

Strevens, M. “The Causal and Unification Approaches to Explanation Unified—Causally”. *Noûs*, 38(1), (2004), 154–176.

Travis, C. “On What is Strictly Speaking True”. *Canadian Journal of Philosophy*, 15(2), (1985), 187–229.

—. *Unshadowed Thought* (Cambridge, MA: Harvard UP, 2000).

Vogel, J. “Tracking, Closure, and Inducive Knowledge”. In S. Luper-Foy, ed., *The Possibility of Knowledge*, 197–215 (Totowa: Rowman Littlefield, 1987).

—. “Subjunctivitis”. *Philosophical Studies*, 134(1), (2007), 73–88.

Woodward, J. “There is no such Thing as a *ceteris paribus* Law”. *Erkenntnis*, 57(3), (2002), 303–328.