Synthese
DOI 10.1007/s11229-015-0841-y

S.I.: CAUSATION IN METAPHYSICS

Natural kinds as nodes in causal networks

Muhammad Ali Khalidi[1]

Received: 12 December 2014 / Accepted: 23 July 2015

Abstract In this paper I offer a unified causal account of natural kinds. Using as a starting point the widely held view that natural kind terms or predicates are projectible, I argue that the ontological bases of their projectibility are the causal properties and relations associated with the natural kinds themselves. Natural kinds are not just concatenations of properties but ordered hierarchies of properties, whose instances are related to one another as causes and effects in recurrent causal processes. The resulting account of natural kinds as clusters of core causal properties that give rise to clusters of derivative properties enables us to distinguish genuine natural kinds from non-natural kinds. For instance, it enables us to say why some of the purely conventional categories derived from the social domain do not correspond to natural kinds, though other social categories may.

Keywords Natural kinds · Causation · Metaphysics of science

1 Introduction

Philosophers from Mill to Millikan think that natural kinds have something to do with causation, but the precise connection is not always made explicit. Another common view, most prominently articulated by Boyd, has it that each natural kind is associated with a loose set or cluster of properties. Putting these two views together, I will argue that rather than an undifferentiated set of properties, $\{P_1, P_2, \ldots, P_n\}$, associated with a natural kind K, the properties associated with each natural kind are causally structured in certain ways. But contrary to Boyd's account, there is not always a causal

✉ Muhammad Ali Khalidi
khalidi@yorku.ca

1 Department of Philosophy, York University, 4700 Keele Street, Toronto, ON M3J 1P3, Canada

Published online: 02 August 2015

mechanism that keeps these properties in homeostasis. This gives rise to a picture according to which natural kinds are identified with one or more causal properties, which when instantiated or co-instantiated, cause the instantiation of other properties in structured causal networks. The causal links between these properties may not be strict and this gives rise to natural kinds with fuzzy boundaries. Moreover, the combinations of properties that cause the instantiation of another set of properties can also vary, leading to natural kinds identified with loose sets or clusters of properties. I will argue that this picture applies just as much to uncontroversial natural kinds like the chemical elements as it does to more contentious kinds in the special sciences, including the social sciences. It also enables us to distinguish genuine natural kinds from clusters of properties that are conventionally rather than causally related, a feature that applies to some though not all kinds in the social domain.

In Sect. 2, I will spell out what I take to be the connection between projectibility and the terms or categories that correspond to natural kinds. Then, in Sect. 3, I will try to show that the projectibility of natural kind terms is grounded in causal relations among the properties that pertain to natural kinds themselves. In Sect. 4, I will take a critical look at three other relationships that are said to hold among the properties (or property instances) associated with natural kinds: coexistence, copying, and convention. Finally, in Sect. 5, I will extend the account to natural kinds based in causal history rather than causal power. The aim will be to provide a unified causal account of natural kinds.

2 Projectibility and natural kind terms

Projectibility is perhaps the most widely attested feature of the categories, terms, or predicates that correspond to natural kinds.[1] When it comes to natural kind terms, what seems to set them apart from other general terms is the fact that they are particularly efficacious when it comes to framing inductive inferences and they feature in many empirically verifiable generalizations. In the absence of a widely accepted definition of projectibility (and bearing in mind that some philosophers think of it as pertaining to hypotheses or statements rather than terms or predicates), I will operate with the following rough-and-ready characterization: predicate P is projectible relative to predicate Q if and only if we can draw a legitimate inductive inference from x is P to x is Q, where the predicates stand for properties or natural kinds and x denotes a particular, whether a concrete individual, a specific event, a particular process, and so on.[2] There are two things to notice about this proposed explication of projectibility. First, it is relativized to pairs of predicates rather than applicable to single predicates. Although we can project from 'x is an emerald' to 'x is green', we cannot project from 'x is an emerald' to 'x is oval' (not to mention 'x is grue'). Of course, we could say that P is projectible (*tout court*) if and only if there is at least one other predi-

[1] But for a recent dissenting opinion, see Ereshefsky and Reydon (2014), who argue that some scientific kind terms are not projectible.

[2] Quine (1969, p. 115) puts it in terms of confirmation rather than inductive inference: "projectible predicates are predicates ζ and η whose shared instances all do count, for whatever reason, toward confirmation of $\ulcorner$ All ζ are η $\urcorner$." I prefer my formulation partly because I think that universal generalizations are vanishingly rare in the sciences.

cate relative to which it is projectible. Second, this characterization does not provide us with a very informative explication of projectibility, since I am presupposing the notion of legitimate inductive inference. Hence, there is no pretense of solving the problem of induction, or saying why grue-like predicates are not projectible, or why ‘emerald’ is projectible relative to ‘green’ but not ‘grue’ (Goodman 1955/1983). But the aim here is not to explicate projectibility and inductive inference, but merely to clarify the connection between natural kind predicates and projectibility (and hence inductive inference).

If we accept the claim that all natural kind terms are projectible (though perhaps not all projectible terms are natural kind terms[3]), we can go on to consider some of the ways in which this projectibility is manifested. Think of every philosopher’s favorite examples of natural kind terms, ‘gold,’ ‘water,’ and ‘tiger,’ and contrast them with terms that do not correspond to natural kinds, say, ‘dirt,’ ‘juice,’ and ‘pet.’ If we compare the types of generalizations and inferences that can be made using the former set of terms as opposed to the latter, the contrast readily emerges. There are many empirical generalizations that can be made using the term ‘gold’ (or the predicate ‘is a sample of gold’), and one can deploy these generalizations to make inferences and to project from one sample of gold to another. Here are a few such general assertions and inferences:

All gold has a melting point of 1337 K.
If x is a sample of gold, then x has a density of 19.3 g cm^{-3}.
If a sample of a substance has a melting point of 1337 K and a density of 19.3 g cm^{-3}, then it is (probably) a sample of gold.

By contrast, there is very little that can be asserted of dirt, since it is highly context-specific what counts as dirt, what is dirt in one context may not be in another, and indeed what makes something dirt in one context may be quite different from what makes it dirt in another. Moreover, different samples of dirt have very little in common and can come in different varieties that share no discernible properties (e.g. sand, oil, soot), neither at the micro-level nor at the macro-level. We cannot even say that all dirt is solid at room temperature, or that all dirt can be eliminated with soap, or that all dirt is airborne. Perhaps the most that can be said about dirt is something along the lines of: all dirt is unwanted material. But that is arguably part of the meaning of the term ‘dirt’, or the property by which we identify it in the first place. Similar points can be made for other terms that do not correspond to natural kinds.

Based on this contrast between an uncontroversial natural kind term and a term that does not pick out a natural kind, it appears that there are at least two dimensions of strength that characterize the projectibility of natural kind terms. First, the paradigm cases of natural kind terms feature in generalizations that are, if not universal, at least stable across contexts and in different circumstances. While these generalizations may contain significant ceteris paribus clauses, they do not have to be hedged to the point

[3] Examples have been put forward of projectible terms that are not natural kind terms. For example, Godfrey-Smith (2011) has argued that certain types of statistical projections do not involve an assumption of “naturalness”. I only claim that projectibility is a necessary condition on natural kind terms, not that it is a sufficient condition.

that they are virtually useless or defeat any attempt to make inferences or projections. For example, some of the generalizations that we have already encountered involving gold need to be qualified in various ways. The statement that all samples of gold have a density of 19.3 g cm^{-3} is not unrestricted; it obtains under conditions of “standard temperature and pressure.” But this does nothing to blur the distinction between terms like ‘gold’ and terms like ‘dirt’, not least because the qualification is itself quite precise and the reasons for making it are well understood (sometimes thanks to results in other branches of science). Still, it does introduce an element of dimensionality into the projectibility of natural kind terms, since some generalizations that can be made using such terms have fewer exceptions or a more limited range of such exceptions than others, though it seems difficult to quantify this difference. That means that projections to new samples are seldom assured and can fail in certain contexts or in some circumstances.

Second, natural kind terms seem to be associated with a wide range of generalizations rather than just one or a few. We can assert many general things about gold, and the natural kind *gold* is associated with a variety of different properties. If I know that the sample of the substance before me is a sample of gold, I can infer that it has a certain melting point and hardness, as well as a certain electrical conductivity, ductility, and reactivity with various other substances. By contrast, terms that do not stand for natural kinds enter into no such generalizations, or at best very few, as we have already seen. But this claim leads to a point of contention among philosophers. Famously, Mill (1843/1882, I vii 4) held that the properties associated with kinds must be “inexhaustible” or “indefinite” in number. However, some nineteenth-century critics of Mill, notably Peirce, objected that this was too stringent a requirement, since many real kinds are not associated with an unlimited number of properties. Indeed, when he attempts to improve on Mill’s definition of “real kind,” Peirce (1901) proposes the following: “Any class which, in addition to its defining character, has another that is of permanent interest and is common and peculiar to its members, is destined to be conserved in that ultimate conception of the universe at which we aim, and is accordingly to be called ‘real.”’ Peirce is surely correct to take issue with Mill’s insistence on inexhaustibility (for further justification, see Khalidi 2013, pp. 51–52), but whether he is right to go to the other extreme, in appearing to allow just one additional property associated with a kind, apart from its “defining character,” is debatable.[4] There may be no principled reason for thinking that there has to be more than one additional property associated with a natural kind, but many of the familiar examples involve considerably more than one. Millikan (1999, p. 48) considers this the “traditional” view of natural kinds:

> … a natural kind corresponds not just to a projectible predicate, but must figure as the subject of many empirical generalizations. No science consists of a single generalization, nor of a heap of generalizations about different kinds of things. A science begins only when, at minimum, a number of generalizations can be made over instances of a single kind…

[4] I don’t think that Peirce’s talk of “defining character” requires a commitment to analytic truth. We can think of the defining characters of a natural kind as those properties used to identify the kind in the first place (see Sect. 3 for more on this).

Whether or not Millikan is right about this being the traditional view of natural kinds, this consideration introduces another dimensional aspect to the projectibility of natural kind terms, whereby some of them are projectible in numerous different ways or relative to a number of other terms, while others are only projectible relative to a few (perhaps as few as one, if Peirce is right).

Therefore, there are at least two dimensions along which one can plot degrees of projectibility for natural kind terms, roughly in terms of the generality of the projections and their variety, respectively (cf. Millikan 2000, p. 26). The two dimensions would seem to be orthogonal, though strength along one of these dimensions may help compensate for weakness on the other, when it comes to their utility for scientific inquiry. A natural kind predicate that has a multitude of other predicates associated with it might be useful in scientific inquiry even if these associations are not perfectly strict or exceptionless, and a natural kind predicate associated invariably or universally with certain other predicates may get away with not being associated with a great variety of such predicates. (In the following section, I will argue that these two dimensions of epistemic strength correspond to two ontological dimensions.)

To sum up, the projectibility of natural kind terms consists in the fact that when it comes to a natural kind predicate K, there is no shortage of other predicates, P_1, P_2, …, P_n, and so on, such that we can reliably assert that if x is K, then x is P_1, x is P_2, …, x is P_n, and we can do so with a high degree of generality. Of course, this presupposes that the predicate K itself has not been introduced to stand in for the conjunction of P_1, P_2, …, P_n; otherwise our assertion would be vacuous. So we are assuming that K is shorthand for some other set of predicates, say Q_1, Q_2, …, Q_m. Hence, what makes natural kind predicates so useful for inductive inference are the robust correlations between them and other predicates. This is what distinguishes terms like 'gold', 'water', and 'tiger' from terms like 'dirt', 'juice', and 'pet'. But this way of broaching the topic, via the projectibility of natural kind predicates, leaves out an important feature of natural kinds. We may well be warranted in projecting from one set of predicates to another, for example from 'x is yellow, shiny, and malleable' to 'x has a melting point of 1337 K' and 'x has a density of 19.3 g cm^{-3}'. (At least, that is the case if we refine the properties that we associate with gold to indicate a particular shade of yellow, a certain lustre, and degree of malleability, so that we do not confuse gold with pyrites and make unwarranted projections.) But even though these projections may be inductively warranted, they leave out a central feature of scientific inquiry. There is clearly nothing special about the complex predicate 'x is yellow, shiny, and malleable.' We might as well have inferred in the other direction, from a specific melting point, density, and thermal conductivity, to a certain color, lustre, and malleability, and we would have been warranted in doing so. The properties that we happen to hit upon first are not always the fundamental or "core" properties, which means that even if we are successful in identifying a genuine clustering of properties (rather than a spurious correlation), further investigation is needed to determine what accounts for this clustering of properties. Mere correlation of properties is not enough, since we are ultimately interested in causation.

We now think that the core property of gold consists of having atomic number 79 and we think of the other properties of gold as somehow arising directly from that property. But this is too simplistic. More precisely, the core properties consist of a certain atomic number as well as a certain mass number, since many of the properties of gold depend not just on atomic number but on mass number or the conjunction of the two. Moreover, only a narrow range of mass numbers results in nuclei stable enough to give rise to the other properties of gold (cf. Khalidi 2013, pp. 167–169). There are only a few isotopes of gold that persist long enough to participate in chemical reactions or exhibit many of the other properties associated with gold. In addition, the macro-properties of gold follow only because atoms with these micro-properties *when aggregated in very large numbers and found in certain contexts against certain background conditions*, will tend to give rise causally to certain other properties, such as color, lustre, density, melting point, thermal conductivity, and so on. Some of these macro-properties can be scaled down to the micro-level, such as (some types of) chemical reactivity, while others cannot. Moreover, some *obviously* cannot (e.g. malleability, melting point), since they are macro-properties par excellence, but others (e.g. density) are harder to assess. To the extent that a single atom can be said to have a density (the ratio of its mass to the volume it occupies), it will not correspond to the density of macroscopic samples of large aggregates of atoms of that type, since the density of the aggregate depends also on the spacing between atoms. Very few, if any, properties of a single atom of gold scale up smoothly to macroscopic samples of gold. These complications are not always kept in view when we say that gold is the element with atomic number 79 and that this is the essential property of gold (necessary and sufficient condition for something to be gold), from which all its other properties flow. Rather, bearing in mind all these sundry complications, the primary causal properties of gold include atomic number 79 as well as a disjunction of mass numbers, which give rise in turn to a cluster of other causal properties (e.g. ionization energies, atomic radius, etc.). Further, when atoms with this cluster of properties are aggregated under certain conditions they give rise to another cluster of properties. The cluster of primary causal properties corresponds to what I have been calling the "core" properties of this natural kind, while the others are "derivative" or causally "secondary" (Khalidi 2013). Indeed, there will typically be a hierarchy or a series of cascading layers of properties, especially when it comes to the particularly fertile natural kinds like the chemical elements. In some cases, the predicate or category associated with the kind is used to denote the core properties, while in others it is used to denote the whole panoply of properties. Either way, what this means is that there are no necessary and sufficient conditions; at best, in this case, there is one property whose presence constitutes a necessary condition (atomic number) and a set of other properties (mass numbers), each of which provides a sufficient condition when conjoined with the necessary condition.

The distinction between the core and derivative properties of natural kinds has not been sufficiently discussed by philosophers, and those who have discussed it have done so primarily from the perspective of an essentialist account of natural kinds. Elder (2007) considers that there is an essential property of natural kinds that "controls" the others. He is concerned mainly to use this conception to rule out certain phase sortals (e.g. *tadpole*, *adolescent*) as natural kinds. But in doing so, he assumes that the properties of natural kinds are modally necessary, intrinsic, and both necessary and sufficient for kind membership. The account I am advocating is not an essentialist one and rejects these criteria for natural kindhood. Though I cannot argue here directly against essentialism, suffice it to say that, on the current proposal, the causal properties associated with natural kinds can be extrinsic or functional, and there may be no set of properties that are singly necessary and jointly sufficient for membership in the kind.[5] Some essentialists identify what I have been calling the "core" properties as the essence and the "derivative" properties as "propria", "properties", or "necessary accidents" (Oderberg 2011). As such, essentialists have what Oderberg refers to as a "unity problem" in explaining why essential properties are themselves co-instantiated or why they are co-instantiated with the propria. Since essentialists consider that these properties co-occur as a matter of metaphysical necessity, according to Oderberg (2011, pp. 93–94), they cannot simply say (with non-essentialists) that the coinstantiation of these properties is a matter of causality or brute fact. From a non-essentialist perspective, however, this does not seem to be a problem. Having said that, there are two questions one can raise about the co-instantiation of the properties associated with natural kinds. The first concerns the co-instantiation of the core properties themselves, which seems to be brought about in more than one way, as I will explain in Sect. 4. The second concerns the co-instantiation of the core properties with the derivative properties, which is brought about as a result of causality, as I will try to argue in Sect. 3.

3 Causality and natural kinds

Natural kind *terms* or *predicates* are projectible precisely because natural kinds themselves are implicated in causal processes.[6] There is clearly an inference here from epistemology to metaphysics, in accordance with what Boyd calls the "accommodation thesis." That is the claim that we choose our categories and delimit their boundaries in such a way as to accommodate the causal structure of the world. The fact that natural kind terms are projectible and feature in true inductive generalizations is a reflection of causal structures. Boyd (1999, p. 148) sums up his conception of accommodation as follows:

> … we are able to identify true generalizations in science and in everyday life because we are able to accommodate our inductive practices to the causal factors that sustain them. In order to do this—to frame such projectible generalizations at all—we require a vocabulary… which is itself accommodated to relevant causal structures.

Such pronouncements are liable to lead to an accusation of mixing epistemology with metaphysics. But that objection is misplaced if we adopt a stance of naturalism. Naturalists believe that philosophical theories ought to be constrained by empirical knowledge. In this case, it means that the broad features of natural kinds must conform with what we know about the natural world (including the social world). We infer what

[5] Moreover, the condition of modal necessity is trivial and can be satisfied by natural kinds and non-natural kinds alike. See Khalidi (2013) for criticisms of essentialism about natural kinds.

[6] Numerous authors draw a connection between natural kinds and causality; see e.g. Mill (1843/1882, IV vii 2), Broad (1920, p. 44), Quine (1969, p. 133), Boyd (1991, p. 139), Kitcher (1992, p. 104), and Millikan (2000, p. 18).

natural kinds exist and discern their commonalities after determining which categories play a central role in our knowledge-gathering enterprises.

In Boyd's account, the "causal structures" that undergird natural kind terms and generalizations take a fairly specific form: they consist of clusters of properties kept in homeostasis or equilibrium by a causal mechanism. Each natural kind is a homeostatic property cluster (HPC), a loose combination of properties that recur regularly and are kept in equilibrium by a causal mechanism. But reflection on a wide variety of kinds suggests that these features are not found in all of them. This is perhaps most convincingly demonstrated by the example of chemical elements, such as gold. Hence, various philosophers have found these requirements too stringent, both the idea that there is a causal *mechanism* that keeps the cluster of properties together and that it does so via a process of *homeostasis*. Boyd's account has seemed to some to be tailored to biological kinds, where feedback loops and homeostatic mechanisms are common (e.g. in organisms, ecosystems), but others have argued that it does not apply well to the paradigmatic biological kinds, namely species.[7] At any rate, it is clear from Boyd's work that he does not regard this causal template to apply to all natural kinds, at best those that derive from the "special sciences," such as biology, geology, and psychology. If Boyd's causal model of natural kinds is too restrictive, what metaphysical conclusion can be drawn from our epistemic practices?

Care must be taken in inferring causal connections from projectibility, since (as I mentioned in Sect. 2) we can sometimes project from effect to cause rather than from cause to effect, as well as from one effect to another effect of a common cause. That is why, once we have distinguished gold from pyrites based on macro-properties, we can project from the fact that a sample of a substance is a particular shade of yellow, has a certain shiny lustre, and is highly malleable, that it has a melting point of 1337 K, that it is ductile, or indeed that it consists of atoms with atomic number 79 bonded together in a crystalline structure. These are legitimate inductive inferences, yet they project either from effects to causes, or from and to different effects of a common cause. Nevertheless, in all of these cases, the projections are parasitic on the connections that obtain between causes and effects. When it comes to natural kinds, causal relations among properties (or more properly, property instances) constitute the ontological ground for the projectibility of the corresponding predicates. But there need be no single causal mechanism that leads these properties to be co-instantiated nor need there be any kind of feedback process that ensures that these properties do not depart from an equilibrium state of co-instantiation. So Boyd's account has to be loosened in such a way as to retain the emphasis on causality without the mechanism or the homeostasis.

As I suggested in the previous section, in the paradigmatic instances of natural kinds there is a causal relationship between the core and derivative properties. We can model these relationships by means of directed causal graphs (cf. Woodward 2003, 38ff),

[7] One of the main sources for this criticism is Ereshefsky and Matthen (2005), and one of the principal responses is Wilson et al. (2009). For other examples of natural kinds that do not conform to Boyd's template, see Khalidi (2013). But it should be noted that Boyd (1989, p. 16) sometimes indicates that the homeostatic mechanism may be metaphorical rather than literal. The account I am proposing is very similar to Boyd's when one drops the mechanism and homeostasis.

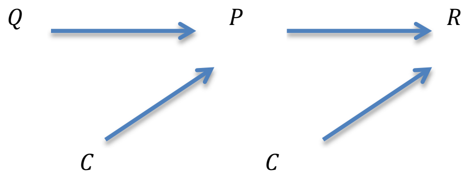

Fig. 1 In this directed causal graph, the core properties of natural kinds ($Q = Q_1, \ldots, Q_n$) cause the derivative properties, whether directly ($P = P_1, \ldots, P_m$) or indirectly ($R = R_1, \ldots, R_k$), in conjunction with background conditions and other properties ($C = C_1, \ldots, C_j$)

which are ordered pairs of edges and vertices, where the vertices represent the relata of the causal relation and the edges represent relations of direct causation.[8] To put it in terms that I have been using so far, the relata include one or more core properties (Q_1, …, Q_n), which can be thought of as variables that take on specific values (whether continuous or discrete, possibly binary), as well as one or more derivative properties (P_1, …, P_m). The derivative properties often lead to the instantiation of yet other derivative properties (R_1, …, R_k) in a causal chain or hierarchy arranged in a temporal sequence (see Fig. 1). They rarely do so on their own, but only in conjunction with other properties (C_1, …, C_j), which can be certain background conditions or the aggregation of such property instances in sufficient numbers (as in the instantiation of the macro-properties of gold). Rather than a strictly linear process, in many cases we observe a web-like network of such causal relationships whereby some causal processes interact with other causal processes (e.g. when gold melts, the properties of the chemical bonds of solid gold interact with the properties of temperature and atmospheric pressure). Eventually, many of these processes bottom out in causal "dead-ends", in the sense that the properties that are instantiated do not tend to consistently generate any further properties even when conjoined with other sets of properties. The networks represented in these directed graphs represent types of causal process rather than tokens, though the graphs are drawn on the basis of token observations of sequences of property instances (e.g. gold melting at a certain temperature). Some vertices may also represent disjunctive combinations of properties (e.g. Q_1 & $[Q_2 \vee Q_3]$). What enables natural kind categories to play the role that they do in our inductive, explanatory, and taxonomic practices is that they consist of highly connected nodes in causal networks. They are represented by those vertices in directed causal graphs from which many edges originate, whether directly or indirectly, leading to other vertices. (I will argue in Sect. 5 that they can occasionally be vertices at which many edges terminate.)

If natural kinds correspond to those (possibly loose) clusters of properties that when co-instantiated, cause the instantiation of a multitude of others (also possibly loosely clustered), then various things follow. The first implication is that natural kinds can have fuzzy rather than sharp boundaries. This follows directly from the loose clustering of properties, especially the core properties or causally prior properties, and this rules out associating each natural kind with a set of properties that are singly

[8] Roughly speaking, *X* is a "direct cause" of *Y* when there is a possible intervention on *X* that will change *Y* when all other variables are held fixed by some intervention, relative to some level of analysis (Woodward 2003, p. 55).

necessary and jointly sufficient for the instantiation of that natural kind. Although some contemporary philosophers treat this as a radical thesis, that may be due to the recent dominance of essentialism in analytic philosophy, since many nineteenth- and early twentieth-century philosophers freely admitted that natural kinds have fuzzy boundaries. Whewell (1840/1847, VIII ii 10) thought that definitions are of no use in delimiting a "Natural Group" of objects and that kinds are given by a type or specimen rather than a set of necessary and sufficient conditions. He also thought that there could be individuals intermediate between two different kinds or subordinate kinds intermediate between superordinate kinds, such as species intermediate between genera, without undermining the reality of those kinds. In an arresting image, Whewell (1840/1847, VIII ii 10) held that "this would not destroy the reality of the generic groups, any more than the scattered trees of the intervening plain prevent our speaking intelligibly of the distinct forests of two separate hills."

Another consequence of this account of the causal structure of natural kinds is that there can be degrees of naturalness, along at least two dimensions. These dimensions are the ontological correlates of the dimensions of projectibility mentioned in the previous section. I have noted that there are some predicates that are strongly projectible (with a high degree of generality) relative to others, while other predicates are only weakly projectible relative to others. Correspondingly, some properties are strictly causally associated with others while others are linked by causal connections that are not as strict. But if causal determinism holds, at least for macroscopic interactions, then it is not literally the case that some causal connections among properties are not strict, but rather that there are sometimes intervening causes that prevent or thwart certain causal connections among property instances from obtaining. For instance, there are few intervening causes that would prevent macroscopic samples of pure gold from having a melting point of 1337 K (e.g. a change of air pressure), but there are many intervening causes that might prevent tigers from engaging in predatory behavior (e.g. a variety of different kinds of changes in the ecosystem). Moreover, some combinations of causal properties, such as those associated with the microstructure of gold, give rise to a multitude of others, while other combinations give rise to a narrower variety of causal properties, such as those associated with electrons. Even though electrons are among the strongest candidates for natural kinds, the properties that they have beyond the three that are used to characterize them in the first place (charge, mass, spin), are limited by comparison to gold. But what they lack on this score, they make up for in having a very strict association between their properties.

This point leads directly to a third consequence, namely that natural kinds are different from natural properties, for at least two reasons. Many natural properties like *mass*, *charge*, *density*, *tensile strength*, and *specific heat capacity* (not to mention more controversial examples, like *fitness*, *rate of inflation*, and *gross domestic product*) are determinables rather than determinates. This means that they range over a spectrum of values rather than take on specific values, and there are very few if any effects that are caused by having some value or another for each of these properties. All massive objects have some causal properties *qua* massive, and all charged particles have some (different) causal properties *qua* charged, but they are relatively limited and generic. Moreover, even when these determinable properties take on specific values and become determinate, they rarely have many effects if not conjoined with certain other properties. All objects with a mass of exactly 1 kg have few properties as a result of having a *mass of* 1 *kg* (e.g. the property of *moving with an acceleration of* 1 $m\ s^{-2}$ *when subject to a force of* 1 *N*). Rather, it is usually specific stable combinations of some set of (determinate) properties that have a rich set of effects, giving rise causally to the instantiation of a multitude of other properties. There is nothing to prevent a natural kind from corresponding to a single property when that property leads to the instantiation of a multitude of others, it just does not seem to be a very common feature of the universe. It is tempting to think of the chemical elements as being instances of this phenomenon, with the single property being determinate values of the determinable *atomic number*. But I have already argued that this does not obtain for the element gold, since most of the causal properties of gold atoms and of macroscopic samples of gold depend not just on atomic number but also on mass number. It is the combination of these two properties (which are not entirely independent, since only some combinations of atomic number and mass number are allowable) that results in gold atoms producing a range of macroscopic effects (when suitably aggregated under certain conditions). This point generalizes to most other chemical elements.

4 Coexistence, copying, and convention

In the previous section, I argued that the ontological ground of projectibility is causation, that properties cluster due to causal processes, and that these property clusters themselves issue in causal processes. Thus, the relationship between core properties and derivative properties is causal, and the reason that the core properties cluster in the first place also appears to be causal in many, though not all, instances.[9] But at least some philosophers have argued in various contexts that there may be other grounds for the clustering or co-instantiation of the core properties of natural kinds, namely: (1) coexistence, (2) copying, and (3) convention. In this section, I will discuss each of these in turn.

The possibility that some properties cluster as a matter of brute fact seems to have been first discussed by Mill (1843/1882). In some sections of his *System of Logic*, Mill points out that while causation often accounts for the coinstantiation of properties associated with natural kinds, in some cases, the properties of kinds are linked by "uniformities of coexistence." He goes on to suggest that uniformities of coexistence, unlike causal uniformities ("uniformities of succession"), only account for the co-instantiation of the *ultimate* properties in nature, "those properties which are the causes of all phenomena, but are not themselves caused by any phenomenon…" (Mill 1843/1882, III xxii 2). Russell, who was notoriously skeptical about the importance of causation for science, seems to have had a correspondingly low regard for natural kinds, holding that "functional laws of correlation" are "probably more fundamental than natural kinds," presumably because the latter are mainly causally grounded and

[9] Is it obvious that the link between the core and derivative properties is always causal? On this account, it follows directly from the fact that I have characterized the relationship between core and derivative properties in terms of causal priority. Others (e.g. Hawley and Bird 2011) seem to characterize it in terms of natural law, but I would argue that laws concerning natural kinds are causal laws. Yet others may do so in terms of metaphysical necessity (e.g. Oderberg 2011) but this presupposes an essentialist view that I reject.

causality is not a fundamental feature of the universe (1948/2009, p. 390). Moreover, he thought the most fundamental kinds (according to the physics of his day), electrons, positrons, neutrons, and protons, had their properties correlated for no known reason. Despite the fact that we now know that, for example, protons have the properties that they do due to the properties of their constituents (quarks), we now think that the core properties of the quarks themselves (charge, mass, and other quantum numbers), as well as the core properties of electrons and other leptons, are co-instantiated as a matter of brute fact. There may yet be entities, such as superstrings, that explain why the elementary particles of the Standard Model have the properties that they do. But ultimately, it seems as though some combinations of properties will just be basic, and the most fundamental natural kinds will possess their core properties as a matter of brute fact (cf. Chakravartty 2007, p. 171). At some point, there may be no further causal story to be told concerning the co-instantiation of certain properties, just a constant conjunction of two or more properties that are always (or nearly always) associated in nature. However, there is no reason to think that this will hold of anything but the most fundamental entities in the universe.

It may be true that in other domains, even beyond the most fundamental, we may not have a full account of why the core properties of natural kinds cluster. But this does not mean that they do not do so due to causal factors, just that we may not know what those causal factors are. In particular, when it comes to certain macro-domains, we often do not have a full explanation of the clustering of macro-properties in terms of micro-properties (including their structures, arrangements, interactions, and so on), but it nevertheless seems as though there are causal connections that lead to the instantiation of those very properties.

Another basis for the clustering of the core properties of natural kinds is supposed to be the presence of a copying process. Millikan has drawn attention to what she calls "copied kinds", kinds whose members share properties because they have been produced by means of a process of copying or reproduction, whereby members are produced from one another or from a common template. In particular, Millikan (2005, pp. 307–308) associates three features with copied kinds: (1) all members have been produced from one another or from the same models; (2) members have been produced by, in, or in response to, the same ongoing historical environment (e.g. other copied kinds); (3) some "function" is served by members of the kind, where "function" is roughly an effect raising the probability that its cause will be reproduced. She also indicates that (1) is the primary characteristic of copied kinds, while (2) and (3) support it. Although the process described by (1) is primary, one could say that members of copied kinds are copied, not by happenstance but because of the other two factors, i.e. environmental pressures and the fact that one being produced raises the probability that another is produced. Millikan thinks that copied kinds constitute a genuinely different type of natural kind, and contrasts them with the more familiar natural kinds, which she dubs "eternal kinds." Members of copied kinds are similar, or share a cluster of core properties, not because of natural laws (like eternal kinds), but because they are copied. She identifies such copied kinds in the biological domain (e.g. *dog*) and social domain (e.g. *doctor*), as well as among artifacts (e.g. *1969 Plymouth Valiant*).

While it is instructive to distinguish copied kinds from eternal kinds, the differences seem overblown. Although causality is not given prominence in Millikan's description of copied kinds, copying is undeniably a causal process whereby a certain set of properties is reproduced (either exactly or approximately) in different individuals. This is achieved primarily by asexual and sexual reproduction in the biological realm, though it can also occur through imitation, cultural transmission, and mass production in the social and artifactual realms.[10] Nevertheless, there is a causal explanation for the co-instantiation of these properties, which alludes either to proximal or distal causes. In either case, the copying occurs for a reason. As stressed in factors (2) and (3) identified by Millikan, the properties that co-occur in members of a single copied kind are there for some reason or because they perform a certain function. To focus first on the biological case, insofar as members of a species share certain synchronic properties they do so primarily because of descent from a common ancestor, but also because those very properties have been acquired as a result of environmental pressures, or are largely adaptive. Similarly, when it comes to a social kind like *doctor*, the reason that doctors share certain core properties, characteristics, and abilities is generally a result of being derived from a common template or as a result of cultural imitation (e.g. adoption of a common curriculum, set of techniques and instruments, Hippocratic Oath). But it is also because those very characteristics and abilities have been shaped to some extent by the needs of their societies and patients, and/or because they have proven successful and achieved their purposes. When those features cease to fulfill their purpose, innovations sometimes occur and certain properties are lost while others are gained. In either case, the process is fundamentally a causal one, though it is also characterized by positive feedback loops.[11] Moreover, some processes outside the biological and social domains have characteristics similar to the copying processes that Millikan describes. Though they may not feature the same robust feedback loops, some chemical kinds may qualify as copied kinds. Consider individual DNA molecules, which have a certain chemical composition as a result partly of chemical laws that dictate that that particular molecular structure is stable. However, they can also be said to have the properties that they do because they have been produced from the same template. They therefore represent an intermediate case between "eternal kinds" and "copied kinds", rendering the distinction rather less strict than might have appeared.

The case of at least some social and artifactual copied kinds may seem to be importantly different from that of biological copied kinds. That is because they raise the possibility of what one might call "pure" copied kinds, members of which are reproduced not because they perform a function or in response to certain environmental pressures. Though it is difficult to think of a clear-cut real-world example, hypothetical ones are not hard to come by. Consider a factory that manufactures certain widgets that are mass-produced at the whim of a capricious billionaire. All the widgets are virtually identical, possessing a large number of properties in common, say $P_1, \ldots, P_n$, and they are all produced on the basis of a common template or blueprint. But

[10] It is interesting that some social kinds appear to be copied (e.g. *parliament*, *confederation*, *trade union*, *corporation*), while others may not be (e.g. *marriage*, *government*, *ritual*, *money*). The latter seem to have arisen independently in different human societies.

[11] The existence of causal feedback loops means that the directed causal graphs described in the previous section will need to be modified to represent the fact that an effect can sometimes causally produce an iteration of the cause.

their properties are not connected in any obvious way to each other and they do not, when co-instantiated, lead reliably to a number of other properties. That is because the widgets are not made in response to people's needs and they do not satisfy any purpose or perform any discernible function. They are just curious artifacts with a number of properties that are associated purely by fiat or decree. If we assume for the sake of argument that some artifactual kinds are at least in principle capable of being natural kinds, this particular type of artifact is not a plausible candidate for being a natural kind. What sets this artifactual kind apart from clear-cut cases of natural kinds is that its properties are not causally structured in any way. Unlike the paradigmatic cases of natural kinds that we have already considered, in which some subset of causal properties gives rise causally to another subset of properties, these properties are not causally linked.

One might protest here that the properties $P_1, \ldots, P_n$, are themselves surely causal properties, such as mass, size, composition, structure, and so on. Indeed, it may be said that one could project from the fact that something is a widget to the fact that it has a mass of precisely 1 kg, that it is 10 cm long, made of plastic, and so on. But what it is to be a widget is just to have the properties specified in the blueprint, so it is not as though we have an independent specification of widgets in terms of a set of other properties. (Recall that in Sect. 2, I argued that when projecting from being a member of kind K to having properties $P_1, \ldots, P_n$, we need to have an independent understanding of what it is to be a certain kind in terms of a number of other properties, $Q_1, \ldots, Q_m$, on pain of vacuity.) Still, the objector might insist, one can project from some of these properties to others, for example, from the fact that it is 1 kg and 10 cm long to the fact that it is made of plastic. Now part of the problem in this case is that the properties mentioned are hardly unique or unusual, which makes the projections precarious at best. However, if we imagined a slight variant on the example, in which the widget had a mass of precisely 1.123 kg, a length of 10.456 cm, a unique shape, and so on, then we would perhaps be warranted in projecting from some of these properties to others. Nevertheless, the causal graph representation of this artifact would be rather minimal, even compared to other artifacts. The initial vertex would consist of the properties of the blueprint, which would be linked by separate edges to vertices representing each of the properties mentioned. But there would be no edges linking any of these properties to each other or jointly to other properties. By contrast with other artifacts, the causal profile of the widgets is impoverished. That does not mean that all kinds must have a function, but the causal profile of artifactual kinds consists largely of their causal functions, so functionless artifacts are at best weak candidates for being natural kinds.

This case vindicates the proposal in the previous section that what sets natural kinds apart is that the properties associated with them are causally structured in some way. The widget case also provides further support for Millikan's characterization of copied kinds, not as kinds whose members merely possess a number of properties in common as a result of a copying process, but also as a result of interaction with the environment and the performance of a function. Otherwise, they do not seem to be good candidates for natural kinds (this claim will be further justified in what follows). Of course, it may be that we find a use for these widgets and that they come to occupy an important niche in the social or technological domains. If so, that might lead us to revise our initial judgment about the widgets, and reconsider our dismissal of them as natural kinds. But it is important to clarify that the case of the copied widgets is not typical of either artifactual or social kinds, as should become clear from considering the third reason that is sometimes given for the co-instantiation of certain properties, namely convention.

In the social domain, some categories are associated with a set of properties as a matter of convention, legislation, or institutional decree. Consider social categories like *senior citizen*, *permanent resident*, or *felon*. Roughly speaking, individuals who are classified in one of these categories are so classified because they have properties that are conventionally rather than causally associated. In some jurisdictions a person is a *senior citizen* if and only if one is a citizen and has attained the age of 65, and this is a matter of legislation. This case bears considerable similarity to that of the functionless widgets mentioned above, since it is one in which certain properties are associated together by fiat (though in this case it is usually by legislation rather than decree and the criterion is not entirely arbitrary but based on facts about human longevity). When properties are associated in this manner as a matter of legislation, decree, or convention, they cluster together not as a result of causality. But if that is all there is to a category, then it is likely not to correspond to a natural kind, for roughly the same reason that the widgets do not: these properties are not causally linked and there are no further properties that are caused by them. Now it is sometimes thought that all social categories are conventional in this way, but it is important to acknowledge that though some of them are entirely conventional or have an element of conventionality, many are not purely conventional. There is a difference between the social categories mentioned above (*senior citizen*, *permanent resident*, *felon*) and social categories like *consumer*, *refugee*, or *psychopath*. Unlike the category of *permanent resident*, which in many jurisdictions has precise conditions attached to it that are enacted by legislation, in most jurisdictions there are no necessary and sufficient conditions that a person must satisfy to qualify as a member of the category *refugee*. It has more to do with that person's participation in certain causal processes, or the person's causal history. For instance, refugees are persons escaping hardship or persecution, they flee or migrate from one society to another, they typically face an adjustment period in the host society, they frequently face discrimination, and so on. These kinds of regularities are ones studied and discussed by social scientists, and they have a causal basis. Similarly for a range of other social categories: at least some of the properties associated with each of these kinds are not so associated by fiat or convention. Rather, they fit the pattern of the natural kinds discussed in previous sections, with a set of core properties that lead causally to others.

I have argued that some social categories may be purely conventional and that their properties may be associated by legislative fiat (though the combination of properties may not be entirely arbitrary like the hypothetical widget). But I have also suggested that many social kinds have an important causal dimension and that they may fit the characterization of natural kinds that I articulated in previous sections. Having made a distinction between conventional and causal kinds, and ruled that purely conventional kinds are not candidates for being natural kinds, it should also be emphasized that when it comes to many social kinds, the causal and the conventional can be entangled in certain ways. For instance, one of the social kinds that I mentioned above, *permanent resident*, is legally or conventionally defined in many jurisdictions, yet it can also have a causal dimension and come to play a causal role in social processes (cf. Khalidi 2014). Similar considerations seem to apply to the category *senior citizen*. But the main point is that insofar as some social kinds are *purely* conventional, they are not good candidates for being natural kinds.

In this section, I have canvassed three other reasons, besides causation, for the clustering of properties: coexistence, copying, and convention. I have argued that bare coexistence applies only to the most fundamental domain in nature. I have also argued that copying is largely a matter of causality, so that copied kinds can be natural kinds (though not all are). As for convention, to the extent that it accounts for the clustering of some properties, I have tried to show that the resultant kinds are not natural kinds. Meanwhile, there are many social kinds that are not purely conventional but at least partly causally based and these do fit the characterization of natural kinds that I have elaborated in previous sections. Moreover, many social kinds are introduced on a conventional basis but come to play an important causal role. In the following section, I will describe another variety of kind, which even though it is also based on causality constitutes a genuinely different kind of kind.

5 Etiological kinds

Millikan sometimes refers to copied kinds as “historical kinds” but this terminology can lead to confusion. The reason is that there is a variety of kinds, which can be thought of as historical or etiological, that can be distinguished from the copied kinds discussed in the previous section. What I have in mind are kinds that are individuated based entirely or primarily on their causal history or etiology. In some sciences, such as evolutionary biology, we study processes that unfold over a period of time and we are interested in the provenance of certain individuals, grouping them together in a kind based on their causal history. Though it is rarely the case that such individuals do not also share a number of synchronic causal properties, the primary criterion for categorizing them as members of a kind is etiology. To avoid confusion with Millikan’s copied kinds, I will call such kinds “etiological kinds” (rather than “historical kinds”). Species are perhaps the best known etiological kinds. Many, if not most, biological taxonomists regard species as being individuated primarily (though not exclusively) by their etiology or history of descent. But etiological kinds can also occur in geology (e.g. *sedimentary rock*), planetary astronomy (*meteorite*), cosmology (*cosmic microwave background radiation* or *CMBR*), and other sciences. In each of these cases, the kind in question is individuated with reference mainly to its causal history or etiology.

Copied kinds can be considered a variety of etiological kinds, namely ones whose causal history includes a copying process. Members of a copied kind have all been produced from each other or from a common template, so they share a particular type of causal history. But as we saw in the previous section, Millikan emphasizes that they are also characterized by certain functional properties and have been shaped by their environments. Hence, they typically share numerous synchronic causal properties in addition to sharing a diachronic trajectory. They are at once a variety of etiological kind and a causal kind.

What about pure etiological kinds? Can they be considered an exception to the causal model that I have been outlining in previous sections? Pure etiological kinds are not easy to find in the sciences. Thus, sedimentary, igneous, and metamorphic rocks, in addition to being individuated according to their causal history, also typically share many synchronic properties. The same goes, obviously, for members of most biological species. But there are some species that exhibit extreme polymorphisms or sexual dimorphisms, which better approximate pure etiological kinds. In the case of dimorphic species, even though males and females separately share many synchronic properties, they have very few properties in common. In such cases etiology provides a rationale for grouping individuals together, despite stark variations in phenotype.[12]

If there are kinds that are wholly or largely individuated on the basis of diachronic causal history rather than synchronic causal powers, then they might seem to pose a problem for the view that I have been advocating. I have argued that natural kinds are associated with clusters of properties that give rise causally to a cluster or clusters of other properties. Etiological kinds would not seem to fit this mold. But there are two considerations that mitigate this conclusion. The first is simply that very few if any scientific kinds are purely etiological; at best, they are partly individuated etiologically. Secondly, and more importantly, etiological individuation is also a matter of discerning the causal structure of the world. Even though the point is to trace causal history rather than causal powers, we are still grouping individuals together based on shared causal properties. This point was clearly articulated by Whewell, whose discussion of "natural groups" can be considered the basis for the modern interest in natural kinds. Thus, Whewell (1840/1847, X i 2) comments on what he calls the "paleontological etiological sciences" (which include geology, comparative linguistics, and archeology), as follows: "All these sciences are connected by this bond; that they all endeavour to ascend to a past state, by considering what is the present state of things, and what are the causes of change." Retrodiction is also a form of projection, and etiological kinds are causally-based too, though the causality is backward- rather than forward-looking. In terms of the directed causal graphs that I have appealed to, etiological kinds can be represented by sets of vertices and edges that all converge on a vertex or set of vertices. They are still nodes in causal networks, though they represent terminal rather than initial nodes.

6 Conclusion

In this paper I have tried to offer a unified causal account of natural kinds. Using as a starting point the widely held view that natural kind terms or predicates are projectible, I argued that the ontological bases of their projectibility were the causal properties and relations associated with the natural kinds themselves. Natural kinds are not just concatenations of properties but are ordered hierarchies of properties, whose instances are related to one another as causes and effects in recurrent causal processes. The resulting account of natural kinds as clusters of core causal properties that give

12 For a detailed treatment of a case of this kind, see Magnus (2012) on the anglerfish, though his handling of the case is a little different from mine.

rise to clusters of derivative properties enables us to distinguish genuine natural kinds from non-natural kinds. For instance, it enables us to say why some of the purely conventional categories derived from the social domain do not correspond to natural kinds, though other social categories may.

Acknowledgments This paper is based on a presentation given at a symposium on Causation and Natural Kinds hosted by the Institute for History and Philosophy of Science and Technology (IHPST), Université Paris 1, in March 2014. I am grateful to the organizers, particularly Andrew McFarland, for inviting me, and to all participants for very helpful feedback.

References

Boyd, R. (1989). What realism implies and what it does not. *Dialectica*, *43*, 5–29.

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

Boyd, R. (1999). Homeostasis, species, and higher taxa. In R. A. Wilson (Ed.), *Species: New interdisciplinary essays*. Cambridge, MA: MIT Press.

Broad, C. D. (1920). The relation between induction and probability (part II). *Mind*, *29*, 11–45.

Chakravartty, A. (2007). *A metaphysics for scientific realism: Knowing the unobservable*. Cambridge University Press.

Elder, C. (2007). Realism and the problem of infimae species. *American Philosophical Quarterly*, *44*, 111–127.

Ereshefsky, M., & Matthen, M. (2005). Taxonomy, polymorphism, and history: An introduction to population structure theory. *Philosophy of Science*, *72*, 1–21.

Ereshefsky, M., & Reydon, T. A. C. (2014). Scientific kinds. *Philosophical Studies*, *167*, 1–18.

Goodman, N. (1955/1983). *Fact, fiction, and forecast*. Cambridge, MA: Harvard University Press.

Godfrey-Smith, P. (2011). Induction, samples, and kinds. In J. K. Campbell, M. O'Rourke, & M. H. Slater (Eds.), *Carving nature at its joints: Natural kinds in metaphysics and science*. Cambridge, MA: MIT Press.

Hawley, K., & Bird, A. (2011). What are natural kinds? *Philosophical Perspectives*, 25, 205–221.

Khalidi, M. A. (2013). *Natural categories and human kinds: Classification schemes in the natural and social sciences*. Cambridge: Cambridge University Press.

Khalidi, M. A. (2014). Three kinds of social kinds. *Philosophy and Phenomenological Research*, *90*, 96–112. http://onlinelibrary.wiley.com/doi/10.1111/phpr.12020/abstract.

Kitcher, P. (1992). The naturalists return. *Philosophical Review*, *101*, 53–114.

Magnus, P. D. (2012). *Scientific inquiry and natural kinds: From planets to mallards*. New York: Palgrave Macmillan.

Mill, J. S. (1843/1882). *A system of logic* (8th ed.). New York: Harper & Brothers.

Millikan, R. (1999). Historical kinds and the 'special sciences'. *Philosophical Studies*, *95*, 45–65.

Millikan, R. (2000). *On clear and confused ideas*. Cambridge: Cambridge University Press.

Millikan, R. (2005). Why most concepts aren't categories. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science*. Amsterdam: Elsevier.

Oderberg, D. S. (2011). Essence and properties. *Erkenntnis*, *75*, 85–111.

Peirce, C. S. (1901). Kind. In J. M. Baldwin (Ed.), Dictionary of philosophy and psychology. New York: Macmillan. http://www.jfsowa.com/peirce/baldwin.htm.

Quine, W. V. (1969). Natural kinds. *Ontological relativity and other essays* (pp. 114–138). New York: Columbia University Press.

Russell, B. (1948/2009). *Human knowledge: Its scope and its limits*. New York: Routledge.

Whewell, W. (1840/1847). *The philosophy of the inductive sciences* (2nd ed., Vol. 1). London: John W. Parker.

Wilson, R. A., Barker, M. J., & Brigandt, I. (2009). When traditional essentialism fails: Biological natural kinds. *Philosophical Topics*, *35*, 189–215.

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford: Oxford University Press.