



Homeostatic Property Cluster Theory without Homeostatic Mechanisms: Two Recent Attempts and their Costs

Yukinori Onishi¹ · Davide Serpico²

© The Author(s) 2021

Abstract

The homeostatic property cluster theory (HPC) is widely influential for its ability to account for many natural-kind terms in the life sciences. However, the notion of homeostatic mechanism has never been fully explicated. In 2009, Carl Craver interpreted the notion in the sense articulated in discussions on mechanistic explanation and pointed out that the HPC account equipped with such notion invites interest-relativity. In this paper, we analyze two recent refinements on HPC: one that avoids any reference to the causes of the clustering of properties and one that replaces homeostatic mechanisms with causal networks represented by causal graphs. We argue that the former is too slender to account for some inductive inference in science and the latter, thicker account invites interest-relativity, as the original HPC does. This suggests that human interest will be an un-eliminative part of a satisfactory account of natural kindness. We conclude by discussing the implication of interest-relativity to the naturalness, reality, or objectivity of kinds and indicating an overlooked aspect of natural kinds that requires further studies.

Keywords Natural kinds · Interest-relativity · Homeostatic property clusters · Homeostatic mechanisms · Causal networks · Stable property clusters

Yukinori Onishi, Davide Serpico: Both authors contributed equally to this work.

✉ Davide Serpico
davide.serpico@unige.it

¹ Department of Evolutionary Studies of Biosystems, SOKENDAI (The Graduate University for Advanced Studies), Shonan Village, Hayama, Kanagawa 240-0193, Japan

² Department of Classics, Philosophy and History, University of Genoa, Via Balbi, 2, 16126 Genoa, Italy

1 Introduction

The *homeostatic property cluster theory* (HPC) was initially introduced by Richard Boyd in response to what he takes to be Ian Hacking's (1991) 'challenge' that property-cluster kinds fail to count as natural kinds because, as far as they have fuzzy boundaries, "what puts things into a family is not nature but people in concert" (Boyd 1991, 128).¹ In order to defend property-cluster kinds from this challenge and to find some natural bases for them, Boyd (1991) postulated the existence of a homeostatic mechanism capable of explaining why those properties are statistically associated with each other and shared by the members of a given kind. In this view, homeostatic mechanisms make natural kinds flexible and stable at the same time: on the one hand, they are flexible enough to admit a certain degree of variability in the properties displayed by different members of the kind; on the other hand, they are stable enough to play a role in our epistemic practices, such as induction and prediction. Categorization of property clusters into kinds will be, then, based on those underlying homeostatic mechanisms rather than judgments of "people in concert."

The flexibility and stability of homeostatic mechanisms make them suited to account for entities in the biological domain, where boundaries and membership criteria can be messy. This characteristic has led many scholars to embrace HPC to investigate the reliability of many biological and psychological concepts, such as mental disorders (Kendler et al. 2011), cognition (Buckner 2015), innateness (Khalidi 2016), and intelligence (Serpico 2018). In his 1999 paper (p. 143), Boyd accepted that properties in some HPC kinds can stick together without the involvement of any homeostatic mechanism. However, the reference to homeostatic mechanisms has traditionally constituted the distinctive mark of HPC theories (on this point, see Lipski 2020), intended as a realist response to Hacking's challenge. Thus, in this paper, we follow Carl Craver (2009, 578) and mostly focus on this prevalent interpretation of HPC, according to which properties of natural kinds are held together by a kind-defining, shared (set of) homeostatic mechanism(s).

Despite its importance within the HPC framework, the concept of homeostatic mechanism has never been entirely clarified. According to Boyd himself, a homeostatic mechanism may or may not be 'underlying'; it can derive from a single cause or not; and it may be internal or external to a given system. For instance, as Kendler et al. (2011, 1149) ask, which of the diverse possible causal processes should we emphasize when we construct a psychiatric nosology? Although this issue is not solely conceptual, without a clear notion of homeostatic mechanism HPC would lose an important part of its appeal.

Looking for a deeper understanding of HPC, some scholars have interpreted the notion of homeostatic mechanism as corresponding to the notion of mechanism delineated by Machamer et al. (2000) and developed by the theorists of mechanistic explanation (on this point see Khalidi 2016, 327). Hereafter, we will refer to this notion as *mechanisms in the strict sense* or *Strict Mechanisms*, according to which a mechanism is "a structure performing a function in virtue of its component parts, component operations, and their organization [as the] orchestrated functioning of the mechanism is responsible for one or more phenomena" (Bechtel and Abrahamsen 2005, 43). In Craver's more intuitive definition, "mechanisms are entities and activities organized together such that they do something [...] that the

¹ Hacking does not seem to be explicitly raising this challenge. For instance, he writes "I have nowhere implied that [common nouns that work by family resemblances] are not natural kinds" (1991, 123), although he also conjectures that "a great many family resemblance nouns [...] may properly be called social kinds" (1991, 123).

components could not do on their own” (Craver 2009, 582). For instance, spatial-memory tasks in rats can be analyzed in terms of how the hippocampus forms a ‘map’ of a maze, how this map is ‘encoded’ in hippocampal neural networks, how synaptic changes are realized by the opening of NMDA receptors, and so forth (see Craver 2015).²

By interpreting Boyd’s homeostatic mechanisms as Strict Mechanisms, some scholars have questioned the suitability of HPC for carving nature at its joints. In particular, Carl Craver (2009) developed three challenges against HPC equipped with the notion of Strict Mechanisms (*Strict HPC*, hereafter), according to which the identification of mechanisms is highly dependent on our explanatory practices and, thus, Strict HPC inherits the mark of human interest associated with Strict Mechanism; that is, the resulting taxonomy will be relative to the purpose of inquiry at hand.³ Another challenge to HPC was provided by Matthew Slater (2015), who argued that many putative natural kinds in the special sciences lack anything like a homeostatic mechanism.

These difficulties have led to two sorts of responses from the philosophical community. On the one hand, some scholars, such as Muhammad Ali Khalidi (2018), have replaced the notion of homeostatic mechanisms with other causal notions in order to offer a unified account of natural kinds. On the other hand, scholars such as Slater (2015) have rejected the notion of homeostatic mechanisms as well as any other causal notions in their characterization of natural kindness. We will take Khalidi’s and Slater’s accounts as two alternative, paradigmatic ways to characterize natural kindness: the former account points at the necessity of refining the metaphysical, causal component of the natural kinds theory; the latter, instead, focuses on the epistemic side of the coin, with the aim of clarifying what feature of property clusters enables us to perform inductive inference. Does Khalidi’s improvement on HPC save it from Craver’s challenge of interest-relativity and Slater’s challenge of limited applicability?⁴ If not, does Slater’s metaphysically slender account provide a satisfactory account of inductive practice concerning natural kinds? In this paper, we will analyze these questions and argue that the answer is negative in both cases. These analyses will lead us to the conclusion that at least some aspects of natural kindness require the reference to some sort of underlying process, and that interest-relativity will be an ineliminable part of an account of natural kindness.

² Depending on different accounts, the definition of mechanism can be more or less stringent. Indeed, as Craver (2018) notices, scholars working on mechanisms “have pursued diverse interests and undertaken different, sometimes incompatible, commitments”—often standing back from metaphysical issues. We shall analyze this point in Sect. 2.

³ In this paper, we will focus on this sort of interest-relativity identified by Craver (2009) and discuss whether Khalidi’s (2018) improvement of HPC can avoid Craver’s challenges. We will briefly discuss the relationship between interest-relativity and the naturalness, reality, or objectivity of kinds in the conclusion. It is worth noting that Boyd (1999) accepted some form of interest-relativity (we thank an anonymous reviewer for pointing this out). However, as we noted above, many understand the HPC theory as capable of providing unconventional bases for categorizing property clusters, and thus addressing Hacking’s objection.

⁴ It should be noted that the motivation behind Khalidi’s refinement of original HPC is not the interest-relativity that Craver points out, but rather HPC’s limited generality. Also, in his earlier account, Khalidi explicitly acknowledges pluralism of natural kinds in different domains of inquiry (2013, 219–220). As such, it is not our purpose to criticize his account as failing to achieve what it meant to. Rather, our aim is to examine the exact ways in which interest-relativity arises in his more articulated account that employs causal graphs (Khalidi 2018). This examination is worth it especially because Khalidi does not discuss Craver’s challenges in his 2018 account. Indeed, as we shall argue, employment of causal graphs invites additional contingency in fixing taxonomies. Jantzen (2015) also discusses whether a weaker notion of mechanism can save HPC from one of Craver’s challenges (we thank Jun Otsuka for pointing this out). We shall consider Jantzen’s analysis in Sect. 3.

The argument goes as follows. In Sect. 2, we summarize Craver's and Slater's challenges to HPC and argue that there are good reasons to suspect that those challenges depend on the notion of mechanism that they assume to explicate HPC, namely, the notion of Strict Mechanism. To see whether an alternative notion can save HPC from those challenges, in Sect. 3, we analyze Khalidi's *causal network node account* (CNN)—which replaces the notion of mechanism with causal processes as captured by causal graphs—and argue that CNN faces the same difficulties as Strict HPC despite the employment of a looser notion of mechanism. This leads us, in Sect. 4, to consider an alternative, plausible account of natural kindness, namely, Slater's *stable property cluster account* (SPC). However, as we shall argue, SPC is too slender to account for some inductive inference on natural kinds, for consideration of underlying process is indispensable for the rational reconstruction of such inference. Drawing on these considerations, in Sect. 5 we conclude that at least some aspects of natural kindness require the reference to some sort of underlying process, and that interest-relativity will be an ineliminable part of an account of natural kindness. Although it is often pointed out that relativity to the purpose of inquiry is compatible with the naturalness, reality or objectivity of categories, we shall suggest that treating all interest-relative categories as equally 'natural' insofar as they are parasitic on some objective causal feature will miss pragmatic aspects of the notion of natural kinds.

2 HPC with Strict Mechanisms and Its Difficulties

One of the foremost challenges to HPC comes from Craver (2009). Here, the author draws on the apparent impossibility of making clear-cut, unambiguous distinctions between mechanisms and, thus, of identifying interest-independent categories in nature related to such mechanisms. Overall, Craver examines whether the HPC theory can be grounded in the philosophy of mechanism, and the answer is negative: if the identity of a given HPC kind is based on a Strict Mechanism, then the HPC account is unable to achieve its goal—i.e., identifying purely objective taxonomies.

Within the common interpretation of HPC discussed by Craver, drawing a taxonomy that 'carves nature at its joints' requires an understanding of the relationship between a property cluster and the homeostatic mechanism that realizes it. For instance, if one discovers that a given mental disorder, e.g., schizophrenia, is underpinned by two kinds of mechanisms instead of one, then one should split schizophrenia up into two disorders. By contrast, if one discovers that two different disorders, say, major and minor depression, are underpinned by one kind of mechanism instead of two, then one should lump the two types of depression into one.⁵ In such a way, the refined taxonomy will more accurately reflect the causal structure of the world and will better serve prediction, generalization and, in the case of diseases and mental disorders, prevention and treatment. However, for this taxonomy to be purely objective, there need to be purely objective ways to tell when mechanisms are of the same kind and when they are different, and it is here that, according to Craver, problems come along.

⁵ These two strategies correspond to what McKusick (1969) calls *splitting* and *lumping* strategies, respectively, which are widely adopted in debates on the causal basis of psychological traits (see Craver 2009, 581–582). For recent criticisms to the view that scientific taxonomies are to be revised for maximizing the homogeneity of their projectible properties and underlying mechanisms, see Lipski (2020).

Alongside other scholars (e.g., Glennan 1996; 2002), Craver embraces the view that mechanistic parthood ultimately depends on the explanatory context: parts are always parts with respect to a decomposition framed by reference to some property or activity displayed by the whole they belong to. Furthermore, a mechanism is always a mechanism *for* a given phenomenon that plays some role in our explanatory practices (see Craver 2015). What factors, among the many influencing the behavior of a mechanism, should be included among its components (i.e., what are the boundaries of a mechanism)? What description of a mechanism, from a very detailed to a very schematic one, is the most appropriate? In Craver's view, any choice will depend on the problem at hand. In other words, the appropriate characterization of a mechanism depends on our epistemic interest: "In the absence of any purposes, goals, and forms in nature, there is no principle for dividing the organism into working parts" (Craver 2009, 589).⁶ Craver argues that this applies to kind-underpinning mechanisms as well and, consequently, HPC is doomed to inherit this sort of interest-relativity. Let us call these difficulties the *boundary problem* and the *degree of abstraction problem*, respectively.

In addition to these problems with the identification of mechanisms, Craver points at cases in which there is no one-to-one correspondence between property clusters and the mechanisms realizing them. For example, several etiological mechanisms can lead to a single constitutive mechanism that is directly responsible for a property cluster (e.g., in the case of HIV infection, the same cluster of symptoms can be brought about by different etiological causes). On the other hand, the same etiological mechanism can give rise to different constitutive mechanisms (and, hence, diverse property clusters) depending on the context (e.g., in *tertiary syphilis*). In cases like these, the causal pathway connecting etiological mechanisms, constitutive mechanisms, and property clusters is not linear (see Fig. 1), and the resulting taxonomy will be different depending on which mechanism one attends to. What mechanism underpins a given natural kind, then? If conventional factors are indispensable for settling the issue, the resulting taxonomy will be inevitably affected by conventional factors as well. Let us call this problem the *no one-to-one problem*.

To summarize, Craver identifies three major problems with the HPC account equipped with the notion of Strict Mechanism:

- The Boundary Problem: Indeterminacy regarding factors that should be included among the components of kind-defining mechanisms;
- The Degree of Abstraction Problem: Indeterminacy regarding the degree of abstraction at which kind-defining mechanisms are characterized, from very detailed to very schematic;
- The No One-to-one Problem: Indeterminacy regarding a kind-defining mechanism that arises from non-linear causal pathways between etiological mechanisms, constitutive mechanisms, and property clusters.

In Craver's view, human interests are indispensable for resolving these indeterminacies; thus, taxonomies involve perspectival elements with respect to the sort of mechanism one is interested in. The preference of one over another will depend on the context of inquiry and purpose, e.g., prevention, treatment, or explanation.

⁶ For a discussion on realist and antirealist accounts of mechanisms, see Kaiser (2017). Here, we will not focus on this topic as it concerns the identification of mechanisms per se. We are rather interested in whether and how problems in mechanistic philosophy affect the natural kinds theory.

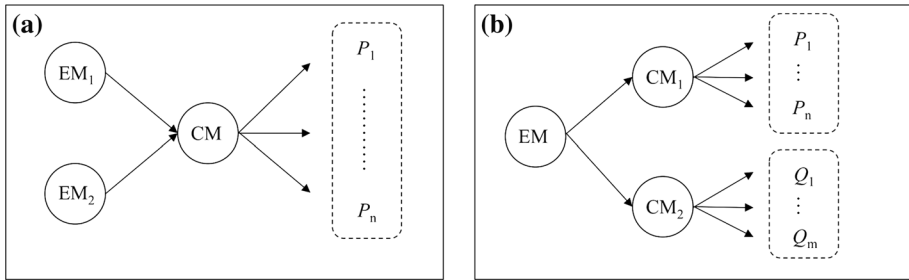


Fig. 1 (a) A case in which different etiological mechanisms lead to the same constitutive mechanism, which gives rise to a property cluster P_1, \dots, P_n (b) A case in which the same etiological mechanism leads to different constitutive mechanisms that give rise to different property clusters

Another challenge to HPC, provided by Slater (2015), is that the notion of homeostatic mechanism is too narrow to account for the natural kindness of many scientific kinds. Thus, Slater says, “Many scientifically important categories—such as elementary particles or chemical species—are associated with clusters of properties whose stability is not plausibly maintained by causal homeostatic mechanisms” (2015, 391). The same applies to biological species: although common selective factors may explain the similarities between the members of a given species, “it is not obvious that these explanations [...] can be interpreted in mechanistic terms” (391).

It is worth noting that the difficulties identified by Craver and Slater could be due to the notion of mechanism the authors assume to explicate HPC, namely, the notion of Strict Mechanism. While these authors are fair enough to assume this notion—especially given that Boyd himself had not offered any detailed account of homeostatic mechanisms—their criticisms may suggest only that the strict notion of mechanism is unsuited for clarifying the HPC theory.

Indeed, there are good reasons to suspect that such a notion is unsuited for this purpose. As we mentioned in Sect. 1, Strict Mechanisms can be described as local, non-aggregative systems (e.g., Craver 2015) or wholes materially composed of *concrete* parts arranged in such a way to produce a specific behavior or function (e.g., Kaiser 2017). However, it has been pointed out that the notion understood in such a strict sense invites some peculiar problems when applied to higher-level aspects of living beings (see DiFrisco 2017; Eronen 2013; Levy 2014; Potochnik 2017). For example, can the relationship between molecular and psychological levels be explained in mechanistic terms? Is an organism a constitutive part of its population in the same sense that a cell is a constitutive part of an organ? Since many questions about the mechanistic framework remain unaddressed (especially with respect to the sorts of higher-level phenomena with which discussions of natural kinds are often concerned, such as psychological traits), it is not reasonable to rely on the notion of Strict Mechanism to explicate the metaphysical basis of natural kinds. In this respect, it is not surprising that HPC faces trouble when the notion of homeostatic mechanism is interpreted as equivalent to the notion of Strict Mechanism.

Slater recognizes this aspect but challenges HPCers to provide an alternative interpretation of homeostatic mechanisms (2015, 393). Craver, in turn, admits the possibility of some alternative interpretation of HPC that is devoid of mechanism, according to which “natural kinds are the kinds appearing in generalizations that correctly describe the causal structure of the world regardless of whether a mechanism explains the clustering of

properties definitive of the kind” (2009, 579). However, he dismisses such an interpretation as “clearly not Boyd’s view” and argues that, in the absence of mechanisms, such an account of natural kinds would be unable to explain “why the kind is so useful for prediction, explanation, and control in a way that conventionalism cannot” (2009, 579). Recently, however, Khalidi (2018) proposed an account of natural kinds that replaces homeostatic mechanisms with a looser causal notion—namely, causal processes characterized by causal graphs. Does the replacement of homeostatic mechanisms with causal processes save HPC from Craver’s and Slater’s challenges? In the next section, we shall turn to this problem.

3 Can a Looser Notion of Mechanism Save HPC?

3.1 The Causal Network Node Account

Khalidi (2018) emphasizes that causal considerations play a central role in our recognition of real kinds in nature. The author reminds us that natural kind terms can allow prediction and induction (in brief, they are projectable) because they can track causal processes. This is arguably what Boyd had originally in mind, too: “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” (Boyd 1999, 148). Notably, however, Khalidi highlights that the notion of homeostatic mechanism should be understood metaphorically rather than literally and proposes a view of natural kinds that is devoid of the literal interpretation of the terms of ‘mechanism’ and ‘homeostasis.’ This looser version of HPC would preserve the emphasis on the causes from which property clusters originate with no need of providing a detailed notion of homeostatic mechanism. In this view, natural kinds are highly connected nodes in causal networks.

Khalidi acknowledges that, “in many cases, we observe a web-like network of causal relationships whereby some causal processes interact with other causal processes” (2018, 1387). For instance, various properties of gold do not derive straightforwardly from the ‘essence’ of gold, but rather are present only if accompanied by other properties, such as temperature and atmospheric pressure. Khalidi holds that these causal networks can be captured by causal graphs, which are ordered pairs of vertices (representing the *relata* of the causal relation) and edges (representing the relations of direct causation).⁷ Then, he argues that natural kind categories “consist of highly connected nodes in causal networks [that are] represented by those vertices in directed causal graphs from which many edges originate, whether directly or indirectly, leading to other vertices” (Khalidi 2018, 1387; see also Woodward 2003). Following Reydon (2015), let us call this the *causal network node* account (CNN, hereafter).

This characterization of natural kinds makes sense of their fuzzy boundaries: that is, the loose clustering of properties avoids any association of natural kinds with property sets that are individually necessary and jointly sufficient for the kind’s instantiation. Moreover, this account appears to be more general than Boyd’s, which, in Khalidi’s view, seems mostly reflective of kind terms employed in “the special sciences, such as biology, geology, and

⁷ More specifically, the *relata* can be one or more core properties or derivative properties. Derivative properties are those caused by core properties, either directly or indirectly, when accompanied with some other properties or background conditions (see Khalidi 2018, Fig. 1).

psychology” (Khalidi 2018, 1386).⁸ Thus, Khalidi’s view arguably extends to most kinds in science, including both uncontroversial natural kinds (e.g., chemical elements) and contentious kinds in the special and social sciences.⁹

Like Boyd’s HPC, Khalidi’s account aims to fulfil two important requirements of a realist theory of natural kinds: first, the reference to causality enables us to discriminate between property sets that are natural kinds and those conventionally tied together; second, the causal relationship between properties (or instances of properties) constitutes the ontological ground for the projectability of the corresponding predicates as well as explanatory and taxonomic practices in science. Is CNN free from Slater’s and Craver’s challenges against Strict HPC, and does it provide a view of natural taxonomies capable of avoiding interest-relativity?

3.2 How Promising Is the Causal Network Node Account?

As we mentioned in Sect. 2, Slater argues that HPC, once equipped with the notion of Strict Mechanism, becomes too restrictive to account for various putative kinds in science. By employing a looser notion of mechanism, CNN seems to be able to bypass Slater’s challenges. As seen in the previous section, CNN is applicable to chemical kinds, such as gold. Likewise, CNN is likely to be able to handle evolutionary cases: indeed, whether natural selection is the cause of evolutionary changes or, rather, the statistical result of some causal process, the general process of phenotypic change itself is a causal process and, hence, it is likely to be captured by causal graphs. Khalidi does not deny that in some cases, the association of properties is a bare fact without any further causal stories to be told, but, he argues, “there is no reason to think that this will hold of anything but the most fundamental entities in the universe” (2018, 1390). For these reasons, Khalidi’s account looks more promising than Strict HPC in terms of general applicability.

Regarding Craver’s challenges, the general question is whether CNN requires perspectival elements in its delineation of natural kinds. In what follows, we will argue that Craver’s three challenges (i.e., the *degree of abstraction problem*, the *boundary problem*, and the *no one-to-one problem*) afflict CNN even though it makes a shift from Strict Mechanisms to the more liberal notion of causal process described by causal graphs.

Let us start by investigating whether CNN can overcome the *degree of abstraction problem*. Considering that this problem mostly arises with respect to the characterization of the components of a mechanism, and given that Khalidi does not need to identify a mechanism in the strict sense, his account may appear to be free from the problem. However, a problem arises with respect to the characterization of causal graphs, for modeling causal networks inevitably involves some degree of abstraction. Jantzen (2015) raises a version of such problem with respect to what he calls *the causal kinds account*, a view that “define[s] natural kinds in terms of a network of causal relations: every system instantiating the same causal structure amongst instances of the same variable types is in the same kind” (2015,

⁸ Note, however, that Boyd (1999) suggests that social structure like feudalism or capitalism, too, might be accounted for by the HPC framework (we thank an anonymous reviewer for pointing this out).

⁹ Remarkably, Khalidi argues for the central role of causal processes even in cases where other interpretations are possible, e.g., the cases of copying, convention, and history (see Millikan 2005). The only kind of entities falling outside Khalidi’s framework are elementary particles or “fundamental entities in the universe,” where the coexistence of properties might represent a “brute fact” (see Khalidi 2018, 1390; see also Chakravartty 2007, 171).

3625–3626). According to Jantzen, “If a kind is characterized by a particular causal structure specified at the finest level of detail [...] then kind induction becomes trivial.[...] [I]t may seem obvious that we need to take a more inclusive view of which causal systems to lump together into kinds. The question is how we can do so without running into Craver’s worry” (2015, 3626).

Considering that CNN seems to be a specific way to implement the causal kinds account in terms of causal graphs, it is likely that the same problem applies here. But how does Craver’s worry concerning Strict HPC transpose to CNN, exactly? Although Khalidi focuses on causal properties and causal processes instead of entities and their activities, entities can be seen as cohesions of properties, and the abstraction of entities means putting aside some of the properties each token has. This possibility suggests that a problem regarding the characterization of a type of entity is likely to arise in its representation with causal graphs, too.

To see this, recall that the original degree of abstraction problem points at the indeterminacy of the appropriate degree of abstraction in characterizing a mechanism, depending on which two instances of mechanisms are regarded as the same type of mechanism (or not) despite having some differences. But, if those instances are different in some features of their components (i.e., the entities or their activities), they must be different in some properties of the entities as well, and, arguably, those differences are present at some levels of abstraction but absent at others. Then, the same question arises: what is the appropriate level at which to characterize a *type* of causal process?¹⁰ Thus, Craver’s challenge regarding the degree of abstraction applies almost straightforwardly to CNN as well. Simply replacing the components of Strict Mechanisms with their properties does not seem to solve the degree of abstraction problem.^{11,12}

What about the *boundary problem*? While the original boundary problem concerns the components that should enter a kind-constitutive mechanism, a similar problem arises in

¹⁰ Note that, according to Khalidi, “[t]he 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” (Khalidi 2018, 1387).

¹¹ One may argue that the appropriate causal graphs should contain all and only those properties that *make some causal differences* and that this constraint will determine the appropriate level of abstraction. However, as Craver objects, “to solve [the degree of abstraction] problem, it will not suffice to demand that the differences in underlying mechanisms must make a causal difference to (or be otherwise explanatorily relevant to) the behavior of a mechanism as a whole because any detectable difference in the underlying mechanism *must make some such causal difference*. Likewise, one cannot object that the difference made is too small or insignificant because such judgments depend on our assessment of which differences are too small to be relevant for our interests, not on the objective features of the *causal structure of the world alone*” (2009, 586; emphasis added).

¹² There are some attempts within the machine learning literature to automatically extract “macro-level” causal structures from “micro-level” experimental data (Chalupka et al. 2016; we thank an anonymous reviewer for bringing this literature to our attention). Although these developments in machine learning might mitigate the degree of abstraction problem in the future, some general remarks are in order. First, even when those algorithms work automatically, some arbitrariness often remains regarding the implementation of algorithms and the choice of hyper parameters, and these choices could be largely affected by pragmatic factors. As Chalupka and colleagues say, “There are many possible alternatives to [Algorithm 1], each with different advantages and disadvantages” (2016, 365). Second, even if those algorithms can automatically extract some causally related features, it is another question whether we can *interpret* them (on the problem of interpretability, see, e.g., Lipton 2018). Thus, taxonomies resulting from such attempts can be very different from ours and it can turn out that many of *our* kind-terms in science are relative to our interest. With these general concerns in mind, it requires careful scrutiny whether the algorithms can provide an interest-independent solution to the degree of abstraction problem.

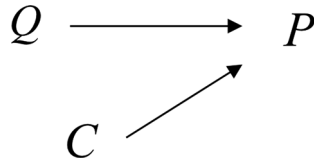


Fig. 2 A directed causal graph that describes the causal process where property Q , together with background condition C , produces property P

Khalidi's account with respect to *the properties (or conditions) that enter a kind-constitutive vertex in the causal graph*. To see this, it is worth noting that there is a sort of indeterminacy with CNN regarding *how to draw a causal graph*.

According to Khalidi, the process that property Q , together with the background condition C , produces property P can be depicted as in Fig. 2. On the other hand, Khalidi also allows that “[s]ome vertices may also represent disjunctive combinations of properties (e.g., Q_1 & [$Q_2 \vee Q_3$])” (2018, 1387). This rule is probably introduced to accommodate cases such as gold, where the core property consists of, according to Khalidi, its atomic number *plus one of the mass numbers* that enable the substance to be stable.

However, this implies that there can be alternative representations of the same causal process, even when we put aside the degree of abstraction problem and attend to the same set of properties. For example, suppose that property Q_1 , together with either Q_2 or Q_3 , can give rise to property P . As shown in Fig. 3, this can be represented in three different ways.

While these three alternative causal graphs may be inter-translatable, the problem is that *the number of vertices in the graph and the number of edges from them may be different in each representation*. This problem will raise a concern for the account of natural kinds that identifies natural kinds with “those vertices in directed causal graphs from which many edges originate” (Khalidi 2018, 1387); for example, it is unclear whether (Q_1 & [$Q_2 \vee Q_3$]) constitutes a single kind or (Q_1 & Q_2) and (Q_1 & Q_3) constitute separate kinds. The whole structure of a causal graph may vary more drastically when more properties are involved in the graph.

This ambiguity leads to the boundary problem that afflicted the Strict HPC. Recall that the boundary problem questions what factors, among those that have *some* influence on the behavior of a mechanism, should be considered as components of the kind-constitutive mechanism. Khalidi points out that various properties that a chemical substance exhibits are present only whereby they are accompanied by other properties, such as background conditions. Such conditions are represented in the causal graph, as in Fig. 2. However, given the alternative ways of drawing causal graphs, why not include background conditions in the vertex and represent it as (Q & C)? Additionally, for cases such as those described in Fig. 3, it is not clear whether Q_2 and Q_3 are part of kind-constitutive vertex or they are merely conditioning properties such that once a kind of objects (as defined by possession of Q_1) possess them, they will exhibit property P .

Consider the case of gold. Khalidi correctly notices that various properties of gold depend not only on its atomic number but also on its mass numbers and that a token of gold can exhibit different properties under different conditions (e.g., different temperatures). Let us suppose that there is a case such that a tiny difference in temperature does not affect its resulting property P . Let Q be a conjunction of gold's atomic number *plus* disjunction of mass

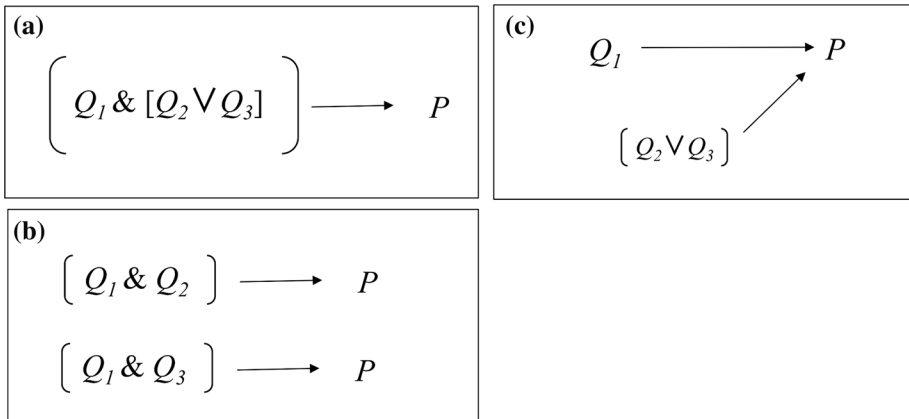


Fig. 3 Three alternative ways to describe the same causal process such that property Q_1 , together with either Q_2 or Q_3 , gives rise to property P

numbers with which gold remains stable,¹³ and let T_1 and T_2 be different values of temperature.¹⁴ When the difference between T_1 and T_2 is negligible with respect to a certain effect P , Q together with either T_1 or T_2 will give rise to property P . Then, the question is whether we should regard Q as a vertex that represents a kind ‘gold,’ whether we should, instead, include a term $(T_1 \vee T_2)$ in the vertex and regard $(Q \& [T_1 \vee T_2])$ as constituting a kind, or whether we should write them separately (i.e., as $(Q \& T_1)$ and $(Q \& T_2)$) and regard each of them (e.g., gold at different temperatures) as being a distinct natural kind. The causal graph account leaves the door open for this sort of ambiguity that arises from the graph-drawing rules, and there is no assurance that such ambiguity will resolve in every case.¹⁵

The proponents of CNN may want to resolve this indeterminacy by offering some reason, either conventional or ‘natural,’ to pick out one of the alternative ways of depicting causal graphs as canonical. For example, CNNers may object that such ambiguity would disappear once further graph-drawing rules are introduced. However, it would be problematic if the number of vertices in a causal graph (and hence the number of natural kinds we identify) varies depending on such a conventional (i.e., interest-relative) matter. Alternatively, CNNers may contend that some properties can be ‘naturally’ regarded as belonging to kind-constitutive vertices while others cannot be. For example, one may argue that kind-constitutive properties should be those that Khalidi calls ‘core properties.’ However, what, then, is the difference between core properties and other properties? Khalidi does not provide an

¹³ That is, Q takes the form of $(A \& [M_1 \vee M_2, \dots \vee M_n])$, where property A stands for gold’s atomic number and properties $M_i (1 \leq i \leq n)$ stand for mass numbers with which gold remains stable.

¹⁴ Khalidi mentions in the explanation of the notation of core properties that they can “take on specific values (whether continuous or discrete, possibly binary)” (2018, 1387). Though this statement concerns core properties, there is no reason to think that this applies to core properties only.

¹⁵ One may object that the problematic case we submit is violating the graph-drawing rule intended by Woodward (2003), which Khalidi (2018) refers to. For Woodward intends that vertices represent variables, rather than their *particular values*, and causal relevance between variables is defined based on whether changes in one of them are conducive to changes in another (2003, 38–45). Thus, one may argue, we are mistaken in treating different values of temperature separately by T_1 and T_2 . However, in Khalidi’s account, vertices seem to represent particular values of variables. For instance, Khalidi himself considers a specific value of atomic number as one of the core properties of gold (2018, 1384).

exact definition of core properties other than that they are “primary causal properties” from which other causal properties (derivative properties) originate (2018, 1384).¹⁶ However, it is far from clear whether we can identify such primary properties in unambiguous ways. For instance, consider the case of spontaneous radio decay, in which Uranium-238 collapses into Thorium-234, popping out an Alpha particle. In this example, whatever property Thorium-234 has is the result of this causal process, which in turn was generated by a property of Uranium-238. Does this mean that the properties of Thorium-234 are all secondary or derivative and, hence, Thorium-234 is not a natural kind? Khalidi discusses various ways of co-instantiation of core properties and admits that “the reason that the core properties cluster in the first place also appears to be causal in many, though not all, instances” (2018, 1389). However, he does not explain why this fact does not strip them of the status of ‘primary’ properties. If what count as primary properties (and hence, the distinction between core properties and derivative properties) depends on the context of inquiry, reliance on this distinction would result in the interest-relativity of natural kinds that Craver contends.

Finally, let us examine the *no one-to-one problem*. Recall that this problem arises when there is no one-to-one relationship between etiological mechanisms and constitutive mechanisms that give rise to property clusters. Sometimes, different etiological mechanisms lead to the same constitutive mechanism (hence, they have the same property cluster); sometimes, depending on the context, the same etiological mechanism results in diverse constitutive mechanisms that give rise to different property clusters. In these cases, the argument goes, the HPC would imply different categorization depending on which mechanism one attends to. Can CNN address this problem?

Khalidi discusses ‘etiological kinds,’ i.e., kinds wherein “the primary criterion for categorizing them as [its] members [...] is etiology” (2018, 1394) and argues that most putative etiological kinds share causal properties as well and can thus be accommodated into his account. However, it is not clear in the first place how some etiological mechanisms (e.g., some type of sexual intercourse or, say, sharing syringes in the case of HIV infection) can be represented as vertices in causal graphs. What properties would the vertices include? How can complex social interactions and intertwined etiological pathways be modeled via causal graphs? There does not seem to be a straightforward answer to these questions.

But let us assume, for the sake of discussion, that etiological mechanisms can be represented as vertices in causal graphs. If so, the problematic cases that give rise to the no one-to-one problem could be represented with causal graphs by replacing mechanisms with vertices. In this case, CNN will advise different categorizations depending on which vertices one attends to in cases where different vertices (corresponding to etiological mechanisms) lead to the same vertex (corresponding to a constitutive mechanism), which gives rise to a property cluster, or in cases where the same vertex leads to various vertices, which gives rise to different property clusters.

It is worth noting that Khalidi himself emphasizes such a web-like manner of property realizations.¹⁷ As such, the no one-to-one problem is not a minor anomaly for CNN. Rather, this suggests the general point that capturing a causal network is one thing and carving out kinds based on causal networks is quite another—likely contextual—matter.

¹⁶ For example, Khalidi says, “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.)” (2018, 1384).

¹⁷ Khalidi says, “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” (2018, 1395).

Note also that Jantzen's concern about the degree of abstraction problem is stated in general terms, leveled at a broad class of causal kinds account. These points suggest that other variants of HPC that employ different characterizations of underlying causal process would face the same difficulty as long as they characterize natural taxonomies as underpinned by a certain part of a causal process, which is characterized at a certain level of abstraction.¹⁸

For the reasons provided in this section, CNN does not seem to save HPC from Craver's challenges and the resulting interest-relativity of natural kinds.

3.3 Does It Help to Employ an Alternative Characterization of Causal Structures?

As we noted above, Jantzen (2015) discusses a causal kinds approach (of which CNN can be understood as a particular formulation), and made the similar point as what we call the *degree of abstraction problem*, alluding to Craver's worry (2015, 3626). As an alternative, he proposes what he calls the *dynamical kinds theory of natural kinds* (DK, henceforth) that retains the emphasis on causal structures but characterizes them in a different, complementary way. Thus, it is worth briefly considering whether this alternative approach could remove the indeterminacies (and the resulting interest-relativity) that afflicts CNN.¹⁹

The central feature of DK is its focus on *dynamical symmetries*, that is, "transformations of a system to which the causal structure of that system is indifferent" (2015, 3761).²⁰ A *dynamical kind* is then defined as "a class of systems of variables that share" those transformations the members are indifferent to (3635).²¹ To take Jantzen's example of chemical analysis, chemists specify constituents of unknown substances by following a decision tree like this: "First add reagent A and check to see if a precipitate forms. If there is a precipitate, follow this or that sub-procedure to identify it. If no precipitate forms, add reagent B, and so on" (3629). In this example, the addition of reagents amounts to transformations to the system (i.e., the solution), and the constituent substance are classified according to their reactivity to those transformations. However, members of the resulting taxonomy would share not only their reactivity to the same transformations, but also their indifference to other transformations in the procedure. The DK approach defines natural kinds based on the latter shared feature: "[possession of the same collection of dynamical symmetries

¹⁸ Note that Anjan Chakravarty, for instance, who also emphasizes the importance of causal process in delineating natural kinds, explicitly admits such interest-relativity (Chakravarty 2007, 175–176).

¹⁹ We thank one of the reviewers for suggesting to discuss this. It should be noted that Jantzen does not explicitly present his account as providing interest-independent bases for categorization; however, it is not unfair to see an attempt in such direction in his account, given that he draws on Craver's (2009) challenge and presents his account partly as a solution to the problem. Also, though DK makes use of formal notions with exact definitions, the following discussion admittedly glosses over many of its carefully developed details for the sake of brevity. Our aim here is only to indicate that it remains to be discussed whether the DK account is free from the problems afflicting CNN.

²⁰ Jantzen explains 'transformations' roughly as "a change of a physical system" (2015, 3631). A typical example in physics is rotation around a certain axis, but he treats as transformations operations such as changing pressure in a container and addition of chemicals to solutions, too (3630, 3634). 'Indifference' in the definition means that, with respect to another operation on some 'index variable' that characterizes the state of the system (e.g., time or concentration of some substance), the final state of the system is not affected by the order in which the transformation and the another operation is applied. See the example of chemical analysis below.

²¹ More precisely, the definition uses another formal notion, *symmetry structures*, and has an additional clause to exclude some trivial kinds of symmetries. See footnote #24.

is] necessary and, to a limited extent, sufficient for induction” (2015, 3636) and thus, he claims, natural kinds are dynamical kinds.

The DK account certainly suggests an important approach to the analysis of natural kinds. However, it is not clear whether a focus on dynamical symmetries instead of causal structures can solve the problems with the causal kinds approach, especially given Jantzen’s remark that the causal and the dynamical symmetry accounts are “two complementary ways” to characterize a system (2015, 3629). Admittedly, Jantzen argues that systems exhibiting different causal structures could share the same dynamical symmetries (i.e., they can be indifferent to the same transformations) and hence, could belong to the same dynamical kind (3636). This would reduce the worry of categorizing every instance of causal structure as a distinct kind, and thus would mitigate the degree of abstraction problem to some extent.

Nevertheless, there seem to remain some indeterminacies concerning the characterization of variables and transformations that serve as the bases for categorization, and it is plausible that human interest is required to resolve such indeterminacies. Notably, the primary examples of Jantzen come from chemistry and physics, or involve very simple biological systems governed by simple equations, such as bacterial colonies. In these contexts, the relevant variables and transformations are relatively clear and often known, e.g., in the form of decision trees. However, it is not clear whether the same procedure is unambiguously applicable to categories in biology, where the transformations and variables characterizing systems are less obvious, and individual differences are abundant.

Take humans, for example, each individual conceived as a system of variables. As is well known, some people exhibit allergic reaction to a certain substance, while others are indifferent to it. If exposure to the allergen counts as a transformation at all, the transformation is a dynamical symmetry for the latter individuals (with respect to some index variables), while it is not for the others.²² That is, the two groups of people do not share a dynamical symmetry and, according to the DK theory, would belong to different kinds.

Although one may find it intuitive that people are categorized based on their different reactivity to a certain substance, given the number of possible transformations that affects people differently, such categorization strategy could end up with a plethora of natural kinds with very small numbers of members. For instance, there are many substances that can cause allergic reactions to some people (but not to others), and some people are allergic to more than one substance. There are also many transformations other than exposure to allergens that people react differently to, reflecting the wide range of individual differences among humans. Recall that DK defines kinds based on the shared set of transformations to which members are indifferent. Then, if we consider the full range of human individual differences (as suggested by personalised medicine) and each of the possible combinations of all the transformations to which people are indifferent, this account would suggest categorizing each individual as a distinct dynamical kind (this would be a plausible consequence of the fact that DK is an essentialist account, see Jantzen 2015, 3626).

A natural way to avoid this would be to focus on certain *types* of allergens and certain *types* of transformations; but, then, the same problem as the degrees of abstraction could arise. Another way out would be to limit our attention to some *relevant* transformations

²² To put this in the way more conform to the definition of dynamical symmetries, for the former ones it makes critical difference whether, for instance, they are exposed to the allergen first and take antihistamine only later or the other way around; on the other hand, for those who are indifferent to the substance, the order does not matter.

(e.g., ignoring difference in people’s reactivity to some transformations), but this clearly invites interest-relativity.

Another indeterminacy with the account concerns what count as transformations and variables that should have relevance to categorizations. For instance, exposure to the same verbal expression (such as an insult) can induce very different reactions in different people. Do such interventions count as transformations? Also, regarding variables, does the presence (or absence) of ‘smile’ after some intervention count as a variable that characterizes the state of a person, and thus affects categorization of humans?²³ If they do, the number of transformations and their possible combinations would skyrocket, and the result would be, again, a plethora of natural kinds including a very limited number of members, if not just one. If they do not, the proponents of DK should provide a natural reason why they should be excluded. The upshot is that a similar problem as the boundary problem could arise in the DK theory, too.

To be fair, Jantzen requires that the collection of the dynamical symmetries that defines a dynamical kind should be *non-trivial*,²⁴ which may mitigate the issue to some extent. However, it remains to be shown exactly how this formal requirement can address the above indeterminacies without appealing to human interest.²⁵

4 HPC Without Underlying Mechanisms: Is It All About Stability?

4.1 The Stable Property Cluster Account

As we mentioned in Sect. 1, instead of loosening the notion of mechanism (like Khalidi), Slater (2015) tries to address Craver’s challenges to HPC by developing an account of natural kindness that drops homeostatic mechanisms. In this account, Slater focuses on the very stability of the property clusters that homeostatic mechanisms are supposed to generate. In other words, Slater focuses on clarifying the feature of property clusters associated with natural kinds that enables us to perform inductive inference; then, he proposes using it as the defining feature of natural kindness.

This approach yields three main benefits (see Slater 2015, 396). First, it avoids the problems with mechanisms discussed in Sect. 2 by focusing on what characteristics natural kinds need to exhibit in order to serve inductive inference on the kinds’ members. Second, the account achieves compelling metaphysical neutrality by avoiding any reference to grounding factors for the stability of property clusters (such as essences and homeostatic mechanisms)—i.e., stability can be realized in many ways. Third, it makes sense of

²³ This is not so unfavorable an example, given that presence or absence of storm is treated as a two-valued variable in Woodward’s (2003) interventionist account of causation, based on which Jantzen develops his account.

²⁴ Roughly put, this requirement is meant to exclude cases such that there are “external causes that influence the states of the system [and] that the variables instantiating the symmetry structure are causally disconnected” (Jantzen 2015, 3635).

²⁵ As regards the no one-to-one problem, in both cases of Fig. 2, DK seems to dictate categorizing property clusters differently depending on their *total causal pathway* (from the etiological mechanism to the cluster). For, when there are differences in the causal pathways leading to the property clusters, there are likely to be some transformations such that one of them is affected and the other is not. If this is the case, the indeterminacy would be resolved, at least for this problem. However, this only comes with the risk of overly fine-grained categorizations, and it is not clear whether such an approach conforms to actual scientific practice.

different kinds of kinds, e.g., essentialist, historical, and HPC kinds. Overall, Slater does not deny the value of searching for underlying homeostatic mechanisms in scientific practice (2015, 402–403), but he believes that this is valuable only in so far as it leads to the discovery of further stable property clusters, on the basis of which natural kinds are ultimately identified (according to the SPC account): “When it comes to the projectability of a kind, the ontological ground—an essence, a causal homeostatic mechanism—is only a means to an epistemically significant end” (Slater 2015, 395).

Then, what sort of property cluster is required for natural kindness? A stable one, of course, but not of any sort. According to Slater, natural kindness requires a sort of ‘cliquish stability’, intuitively defined as follows: “Spotting Peg, Quinn, and Ralph at the mall means that Sarah and Tim [from the same clique] are probably there as well.[...] Call this conception of stability ‘cliquish stability.’ [...] The idea is to capture the fact that some properties are clustered in such a way that possession of some of them reliably (if imperfectly) indicates the possession of the whole cluster (if not each property in the cluster) at that time” (Slater, 2015, 397).²⁶ Thus, Slater claims that natural kinds are characterized by such stable property clusters and calls this view the *stable property cluster* account (SPC).

It should be noted that Slater’s definition of cliquish stability opens the door for some sort of domain-relativity (namely, relativity to the interest of inquiry in a given domain). Indeed, cliquishness is characterized as stability under counterfactual situations, and hence, the question arises of what counterfactual situations should be taken into account. Slater answers that they are those that are relevant to the epistemic practice in question. He also claims that it is necessary to relativize natural kinds to physical contexts as well (he calls this ‘context-relativity’), for the stability of some property clusters can be maintained only under certain physical conditions (such as *in vitro* conditions) (2015, 404). According to Slater, this implies that natural kinds (in the SPC view) are not ontological categories. Instead, he suggests thinking of natural kindness “as a sort of status that things or plurality of things (from various ontological categories) can have” (2015, 406–407). Although this approach certainly invites concerns of relativism, he claims that natural kindness is “understandably treated as a fixed, objective matter when it is highly insensitive to the differences across our classificatory norms and practices” (406–407) and emphasizes that the account does not make natural kinds an arbitrary matter.

Overall, SPC offers a very slender account that provides no straightforward ontological basis for natural kinds. Arguably, this will significantly decrease the appeal of the account in metaphysical terms. Moreover, SPC may not be able to offer an account of the semantics of natural kind terms (on this point, see Häggqvist and Wikforss 2018). Slater recognizes this issue but argues that, “[g]iven the gain of genuine theoretical unity elsewhere, [the problem of semantics] is a potential cost [he is] willing to pay” (2015, 402)—wherein ‘theoretical unity’ means the generality and the neutrality of SPC we mentioned earlier. One may find that these costs are worth paying insofar as the SPC provides an account of inductive inference concerning natural kinds. However, does it provide such an account?

4.2 Does SPC Achieve What It Aims for?

SPC aims to answer the question of “what features must kinds have to be ‘apt’ for inductive inference” (Slater 2015, 385). In other words, it aims to explicate the nature of property

²⁶ For the formal definition of cliquishness, see Slater (2015, 400).

clusters associated with natural kinds *only insofar as it is necessary for accounting for our inductive practice*, putting aside metaphysical questions. However, in what follows, we shall argue that SPC fails to account for some of our inductive practice involving natural kinds: first, inductive inference regarding *previously unknown* properties and, second, inference based on context-dependence of property realization.

Our first concern is that SPC does not provide any account of why it is sometimes reasonable to expect that members of a natural kind that share a number of properties will share other *previously unknown* properties. In the case of ‘thicker’ accounts of natural kinds, such as traditional essentialism and the HPC account, those common groundings offer the reason. If, for example, various properties of the members of a kind originate from a common mechanism, we have reasons to expect that the kinds’ members will also share other properties that originate from such mechanism. In the case of SPC kinds, however, the same reasoning does not apply because the SPC account lacks reference to any underlying mechanism or essence that brings about the shared properties. Thus, when SPCers regard a category as ‘a natural kind’ *solely based on a cliquishly stable cluster of known properties*, this does not give them any reason to expect that the kind’s members have other properties in common apart from those that have been used to define the category.

For example, Burrows and Sutton (2013) found tiny gears on the hindleg trochanter of a planthopper nymph that are used for synchronizing the movement of the legs when the nymph jumps. Having identified such a feature in a nymph, it would be reasonable to expect that it would be found in other nymphs as well because such a complex feature cannot be accidental and there should be some ‘reason’ behind it, be it genetic, selective, or developmental. If the notion of natural kind involves the idea that a kind’s members share such a ‘reason,’ it is reasonable to expect that other members will possess this feature as well. On the other hand, if the notion of natural kind only involves cliquish stability of properties, there is no reason to expect that the new property will be shared by other nymphs.²⁷ The consideration that the discovered feature does not seem to be accidental constitutes an indispensable part of the above reasoning. An account of natural kindness that solely focuses on property clusters is unable to account for this sort of inductive inference in science.

In our understanding, the role of kinds in inductive reasoning is precisely what an account of natural kinds is expected to account for. While there is disagreement among scholars about how many properties natural kinds should share and how strongly cohesive they should be, one of the intuitive differences between natural kinds and arbitrary categorizations is that, while membership in an arbitrary category gives us no reason to expect that its members will share properties other than those used to define the category or those that can be easily inferred from them, membership in a natural kind gives us such a reason (cf. Mill 1846, Vol. 1, Ch. VII §4).

Let us now turn to a second problem with the SPC account. SPC fails to account for another important aspect of inductive inference in science, namely, the context dependence of property realization. As Khalidi emphasizes, it is not that all the properties of gold stem

²⁷ An anonymous referee brought to our attention that cliquish stability at some higher-taxa level, too, might provide such a reason. However, the issue at stake (i.e., there is no principled reason to believe that the newly discovered property belongs to the previously known cliquish property cluster) is likely to hold even when the nymph with the previously unknown property shares a number of properties with other nymphs in the taxon. Indeed, SPC leaves it open whether the newly discovered property is part of the property cluster of the previously known kind (but, for some reason, it was never observed, yet) or rather suggests the existence of a new kind under the same taxon.

from its atomic number, nor is it that all the samples of gold share more or less the same cluster of properties all the time: for instance, its macroscopic properties (such as color, luster, density, melting point, and thermal conductivity) stem from its micro-properties *only when* “*aggregated in very large numbers and found in certain contexts against certain background conditions*” (2018, 1384, original emphasis). Knowledge of these conditional properties is an essential part of successful inductive practice. This is the aspect that Khalidi tries to capture with his causal graph representation, and it is due to this attention to the context-dependent process leading to property clusters that CNN can lump together different property clusters as different states of *the same* kind (e.g., gold), all stemming from the same node (representing the kind) under different conditions. The characterization of natural kindness simply as a cliquish clustering of properties seems insufficient to capture such inference involving natural kinds.

SPCers might object that the context dependence of property realization can be handled simply by relativizing the set of cliquishly stable property clusters to each context. Thus, one may argue, ‘gold’ is a kind such that its members cliquishly exhibit such and such properties under such and such conditions. However, what is missing in this treatment is *the relation among property clusters under different conditions*. For example, chemical substances and organisms possess very different property clusters under different conditions and at different developmental stages. On what bases can SPCers regard them as the *different states of the same kind* (rather than *different kinds*), and how can they account for the inductive inference across different conditions of the same kind? If that relation can also be accommodated to SPC in some way, then it is likely that SPC, despite its appearance, turns out to be very similar to CNN because it would need to attend to processes underpinning property clusters, whether or not one interprets that process as causal or mere propensities.

Chakravartty (2007) clearly sees the importance of such causal relations among properties. He holds that, as a matter of metaphysical facts, certain causal properties are systematically ‘sociable’²⁸ in various degrees, but in addition, these properties also stand in certain relations to other properties, as typically summarized in law statements. According to Chakravartty, it is this sociability of properties *plus* the causal relations among them that “underwrite the inductive success associated with natural kinds in the sciences” (2007, 170).²⁹ The SPC account seems to be attending to only one of these sources of inductive success.

By focusing on explicating what characteristics of natural kinds enable inductive inference, SPC provides an apparently attractive metaphysically slender account of natural kindness. However, the above consideration suggests that reference to some underlying causal process constitutes an indispensable part of such an explication. It is then questionable whether avoidance of reference to underlying causes is a fair cost to pay for a lower metaphysical load.

5 Concluding Remarks

In this paper, we argued that the notion of Strict Mechanism, which is widely employed in mechanistic explanation, is unsuited to clarify Boyd’s notion of homeostatic mechanism and, in turn, the HPC account of natural kinds. Then, we examined two alternative

²⁸ Note that Slater also uses the notion of sociability, but in a different way from Chakravartty. The latter’s notion of sociability is closer to Slater’s notion of cliquishness.

²⁹ Khalidi’s CNN account can be seen as an attempt to develop this picture by employing causal graph representations.

refinements of cluster theories of kinds. The first one is Khalidi's *causal network node* account, which replaces homeostatic mechanisms with causal networks (and its nodes) as captured by causal graphs. We argued that this account faces the same problems as HPC equipped with Strict Mechanism because human interests (purpose of inquiry at hand) play indispensable roles in the description of causal graphs and the delineation of categories out of them. Also, since such indeterminacy is likely to arise in any approach that bases natural kinds on causal processes, we suspect that interest-relativity is an inevitable aspect of natural kinds. The second alternative we assessed is Slater's *stable property cluster* account, which dispenses with the metaphysical discussion on the grounding of property clusters and focuses instead on explicating the features of property clusters that enable inductive inference. We argued that Slater's model cannot account for some inductive inference on natural kinds because consideration of underlying causal process is indispensable for them. If these arguments are correct, the delineation of categories based on causal structures will involve human interest. What does this sort of interest-relativity mean?

First, it means that *there is no ready-made taxonomy on the part of nature until such human interest is specified*. Second, however, as various scholars pointed out, such interest-relativity of categories is compatible with their reality or objectivity (see, e.g., Boyd 1999; Chakravarty 2007; Dupré 1993). While holding a certain purpose is a human matter, whether a given taxonomy is apt for that purpose depends on the way the world is (including its modal structure, however it is explicated). In other words, there are some *constraints posed by nature* on the possible taxonomies that serve for a given purpose.³⁰ Thus, interest-relativity of this sort at most implies *pluralism* and does not fall into relativism. Also, there is still a sense in which those interest-relative categories are 'natural,' i.e., in the sense that *they are parasitic on the causal structure of the world* or that *they conform to constraints of nature* (the way the world is). The 'reality' or 'objectivity' of those kinds could also be explicated as that the features exploited for the categorization (theoretical properties such as 'atomic number' and their supposed causal effects) are the real features of nature (in a correspondence sense). Questions about the reality of a given kind term in science will then become an empirical issue (if we put aside philosophical skepticism) or a matter for the scientific realism debate.

What seems to be missing in the current picture, or seems at least equivocal, is the relative status of those categories each apt for certain purpose. For example, according to Dupré's *promiscuous realism*, "there are many *equally legitimate* ways of dividing the world into kinds" (1993, 6, emphasis added), albeit he also acknowledges that "the naturalness of kinds will be a matter of degree," where the degree depends on the extent to which members of a kind will share a common essence (1993, 63).³¹ Chakravarty, on the other hand, criticizes Dupré on the latter point and argues that "no classification, so long as it is made on the basis of properties in the world [...] is more or less objective than any other," allowing that even classification of objects based on just their mass could be objective natural kinds (2007, 178–179).

³⁰ Here, 'constraint posed by nature' means that the range of taxonomies *suited for a given purpose* is restricted by how the world is. However, it also involves restrictions on possible entities that *can* exist in the first place, i.e., restrictions such that certain entities *cannot exist* (in nature, or for a long time, etc., as is the case for some isotopes or genotypes) because of the way the world is.

³¹ Khalidi also suggests allowing degrees of naturalness along the same line, though he suggests further possible dimensions such as the strength of causal relations among properties (2013, 212–213).

These remarks seem to imply that all interest-relative categories can be treated as equally ‘natural’ insofar as they are parasitic on some objective causal feature. However, allowing too permissive notions of naturalness, reality, or objectivity of interest-relative kinds, and treating every taxonomy equally in this respect, risks to miss the point at stake in some discursive contexts. For instance, it is widely known in Japan that *Echizen crab* and *Matsuba crab*, despite their huge differences in taste and price, belong to the same biological species; namely, they are just ‘brand’ names indicating where those crabs are landed. When people say that the distinction between the two crabs is ‘not real’ or is artificial, they do not mean that the taxonomy is not based on any objective feature, nor that it is not apt for any purpose at all. Rather, what they mean seems to be that the distinction is made for a very particular purpose of human cuisine and the related economic interest. Thus, ‘naturalness’ or ‘reality’ of category (or lack thereof) is here used to mark such contrast in the generality of the purpose those taxonomies are made for.³² In such cases, it would be odd to suggest to treat all interest-relative categories on a par just because they are parasitic on some objective causal features. This would fail to capture what is really at stake in the discourse (i.e., the contrast in generality of purpose).

Likewise, when people discuss whether a given category (such as human race) is natural, what is meant by the question could be whether/to what extent it is due to social constraints or biological ones (in other words, how contingent it is on the societal factor). In such cases, it would be odd to suggest treating any taxonomy of human populations on a par as far as they are based on some causal properties (including those effective in a given social setting). This would fail to capture what is really at stake in the discourse (i.e., social/biological contrast).

These considerations suggest that there are both *epistemic* and *pragmatic* aspects in the notion of natural kinds, and that they are relative to human interests in differing ways. Epistemic aspects concern the ability of categories to support inductive inference, and differing status of categories is, as we argued, characterized by their aptness for various purposes of inquiry, which in turn are parasitic on causal structure—on this aspect, Chakravarty’s account (2007) is very close to our own view. On the other hand, pragmatic aspects concern the function that the notion of natural kinds (or ‘natural,’ ‘objective,’ ‘real’ taxonomies) plays in a given discourse (i.e., what contrast such notions are intended to make in a given discourse, such as those in generality of purpose, or biological/social contrast). For this aspect, it is not clear whether the differing status of categories is characterized by their aptness for certain purposes (as in the case of the epistemic aspect), for there is not always any particular purpose in the discourse (e.g., people do not need any particular purpose to wonder whether there is a real difference between crabs or human races in the sense discussed above). Rather, the differing status of categories in the pragmatic sense would be better captured by reference to the contrast at stake in each discursive context.

Although it is natural that philosophers of science have mostly focused on the epistemic aspect, treating categories on a par based on the analysis of such aspect may miss their difference in the pragmatic aspect, and this may lead philosophers to give somewhat odd

³² This is not to say that biological species concept is free from any human interest, nor that all categories can be linearly ordered in terms of their generality of purpose. Dupré also recognizes the difference in the degree of anthropocentrism involved in scientific taxonomy and ordinary language classifications (1993, 35). Note also that this difference in ‘naturalness’ is not necessarily characterized as the extent to which members share common properties, as Dupré’s (1993) comment in the previous paragraph suggests. For, in this respect, members of the brand-name categories of crabs share more properties (e.g., flavors characteristic of each brand) than when they are put under the same biological category.

adjudications on actual cases in which naturalness, reality, or objectivity of categories are in question. It is beyond the scope of this paper to fully explicate the pragmatic aspect, but a complete account of natural kinds will require further research on such aspects.

Acknowledgments This research was supported by the JSPS KAKENHI (Grant Numbers JP18H00604, JP18K12178), the University of Turin, and the Department of Classics, Philosophy and History of the University of Genoa. One of the authors is grateful to SOKENDAI (The Graduate University for Advanced Studies) for accepting him as Visiting Scholar, which enabled this collaborative research. We thank Matthew Slater, Sören Häggqvist, Yuichi Amitani, Andrea Borghini, and two anonymous reviewers for thoughtful comments and discussions on previous versions of this manuscript.

Funding Open access funding provided by Università degli Studi di Genova within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2), 421–441.
- Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies*, 61(1–2), 127–148.
- Boyd, R. (1999). Homeostasis, species, and higher taxa. In R. A. Wilson (Ed.), *Species: New Interdisciplinary Essays* (pp. 141–185). Cambridge, MA: MIT Press.
- Buckner, C. (2015). A property cluster theory of cognition. *Philosophical Psychology*, 28(3), 307–336.
- Burrows, M., & Sutton, G. (2013). Interacting Gears Synchronize Propulsive Leg Movements in a Jumping Insect. *Science*, 341(6151), 1254–1256.
- Chakravartty, A. (2007). *A Metaphysics for scientific realism: Knowing the unobservable*. Cambridge: Cambridge University Press.
- Chalupka, K., Eberhardt, F., & Perona, P. (2016). Multi-level cause-effect systems. In: *Artificial Intelligence and Statistics* (pp. 361–369).
- Craver, C. F. (2009). Mechanisms and natural kinds. *Philosophical Psychology*, 22(5), 575–594.
- Craver, C. F. (2015). Levels. In T. Metzinger & J. M. Windt (Eds.), *Open MIND*. Frankfurt am Main: MIND Group.
- Craver, C. F. (2018). *Review of Stuart Glennan // The New Mechanical Philosophy*. Retrieved from <https://bjpsbooks.wordpress.com/2018/02/19/stuart-glennan-the-new-mechanical-philosophy>.
- DiFrisco, J. (2017). Time scales and levels of organization. *Erkenntnis*, 82(4), 795–818.
- Dupré, J. (1993). *The disorder of things: Metaphysical foundations of the disunity of science*. Cambridge, MA: Harvard University Press.
- Ereshefsky, M. (2018). Natural Kinds, Mind Independence, and Defeasibility. *Philosophy of Science*, 85(5), 845–856.
- Eronen, M. I. (2013). No levels, no problems: Downward causation in neuroscience. *Philosophy of Science*, 80(5), 1042–1052.
- Glennan, S. (1996). Mechanisms and the nature of causation. *Erkenntnis*, 44(1), 49–71.
- Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of science*, 69(S3), S342–S353.
- Hacking, I. (1991). A tradition of natural kinds. *Philosophical Studies*, 61(1–2), 109–126.
- Häggqvist, S., & Wikforss, Å. (2018). Natural Kinds and Natural Kind Terms: Myth and Reality. *The British Journal for the Philosophy of Science*, 69(4), 911–933.
- Jantzen, B. C. (2015). Projection, symmetry, and natural kinds. *Synthese*, 192(11), 3617–3646.

- Kaiser, M. I. (2017). The Components and Boundaries of Mechanisms. In S. Glennan & P. Illari (Eds.), *The Routledge Handbook of Mechanisms and Mechanical Philosophy* (pp. 116–130). New York: Routledge.
- Kendler, K. S., Zachar, P., & Craver, C. (2011). What kinds of things are psychiatric disorders? *Psychological medicine*, 41(6), 1143–1150.
- Khalidi, M. A. (2013). *Natural Categories and Human Kinds*. Cambridge: Cambridge University Press.
- Khalidi, M. A. (2014). Naturalizing kinds. In B. Bashour & H. D. Muller (Eds.), *Contemporary philosophical naturalism and its implications* (pp. 115–135). New York: Routledge.
- Khalidi, M. A. (2015). Three kinds of social kinds. *Philosophy and Phenomenological Research*, 90(1), 96–112.
- Khalidi, M. A. (2016). Innateness as a natural cognitive kind. *Philosophical Psychology*, 29(3), 319–333.
- Khalidi, M. A. (2018). Natural kinds as nodes in causal networks. *Synthese*, 195(4), 1379–1396.
- Levy, A. (2014). Machine-Likeness and Explanation by Decomposition. *Philosophers' Imprint*, 14(6), 1–15.
- Lipton, Z. C. (2018). The myths of model interpretability. *Queue*, 16(3), 31–57.
- Lipiski, J. (2020). Natural diversity: A neo-essentialist misconstrual of homeostatic property cluster theory in natural kind debates. *Studies in History and Philosophy of Science Part A*.
- Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of science*, 67(1), 1–25.
- McKusick, V. A. (1969). On lumpers and splitters, or the nosology of genetic disease. *Perspectives in biology and medicine*, 12(2), 298–312.
- Mill, J. S. (1846). *A System of Logic*. New York: Harper & Brothers Publishers.
- Millikan, R. (2005). Why most concepts aren't categories. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 305–315). Amsterdam: Elsevier.
- Potochnik, A. (2017). *Idealization and the Aims of Science*. Chicago: University of Chicago Press.
- Reydon, T. A. (2015). Natural kinds no longer are what they never were. *Metascience*, 24(2), 259–264.
- Serpico, D. (2018). What kind of kind is intelligence? *Philosophical Psychology*, 31(2), 232–252.
- Slater, M. H. (2015). Natural kindness. *The British Journal for the Philosophy of Science*, 66(2), 375–411.
- Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford: Oxford University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.