**Special Section** **Open Access**

Rico Hauswald

# The Ontology of Interactive Kinds

DOI 10.1515/jso-2015-0049

**Abstract:** This paper defends the notions of an interactive kind and a looping effect as features of social and human scientific classifications and aims to give a realist interpretation of them. I argue that interactive kinds can best be modeled as a special case of changing causal property cluster kinds. In order to do so, I develop a typology of looping effects according to the sort of entities that are affected, the main types of which are individual-looping, category-looping, and kind-looping. Based on this distinction, I identify interactive kinds as those causal property cluster kinds that are subjected to kind-looping.

**Keywords:** Looping effects; Interactive kinds; Social scientific classifications; Ian Hacking; Causal property cluster theory.

## 1 Introduction

According to an influential argument advanced by Ian Hacking in a series of publications, the kinds studied by the social and human sciences – more precisely, those kinds whose instances are human beings – are "interactive kinds" (see, e.g. Hacking 1995, 1999). As opposed to instances of "indifferent kinds", people can become self-aware as being of a certain kind and/or as being classified in a certain way (for instance, by scientists). As a result, they may change their properties (their behavior, emotions, self-concepts, etc.), which in turn may affect their status as instances of the kind in question, or affect the kind or classification itself: "What was known about people of a kind may become false because people of that kind have changed in virtue of how they have been classified, what they believe about themselves, or because of how they have been treated as so classified. There is a looping effect." (Hacking 1999, p. 104) Classifying people can, for example, take the form of a self-destructive, but also of a self-fulfiling prophecy. Plausible examples, among many, may be found among "transient

**Rico Hauswald,** Department of Philosophy, Dresden University, Zellescher Weg 17, Dresden 01069, Germany, e-mail: rico.hauswald@tu-dresden.de

mental illnesses" (Hacking 1998) or in the context of labeling approaches to criminal behavior (Becker 1963). These and many other objects of social and human scientific research obviously turn out to be "moving targets" (Hacking 2007a): "new knowledge about 'the criminal' or 'the homosexual' becomes known to the people classified, changes the way these individuals behave, and loops back to force changes in the classifications and knowledge about them" (Hacking 1999, p. 105).

It is easy to agree that Hacking's idea has intuitive appeal. Indeed, a number of authors from different disciplines have found it attractive and adopted (part of) his terminology (Hirschfeld 1996; Dupré 2002; Berreby 2005; McNally 2011; Williams 2015). On the other hand, as other authors have pointed out, the notions of an "interactive kind" and a "looping effect" face serious problems. For example, the distinction between interactive and indifferent kinds has been characterized as untenable, or at least in need of revision (Cooper 2004; Khalidi 2010). Another problem is that, from an ontological point of view, Hacking is rather sloppy when it comes to precisely characterizing what goes on in looping effects (Tsou 2007). On the one hand, he suggests that "'[i]nteractive' is a new concept that applies not to people but classifications" (Hacking 1999, p. 103). On the other hand, he states: "I do not mean *only* the self-conscious reaction of a single individual to how she is classified. I mean the consequence of being so classified for the whole class of individuals" (Hacking 1999, p. 115; my emphasis). In these passages, rather different sorts of things are mixed together. Hacking does not make it sufficiently clear what the possible relata of looping effects may be. On the one hand, he speaks of changes of *classes* or *classifications* (that is, linguistic entities, concepts). On the other hand, his very term "interactive *kind*" suggests that he means kinds in the strict sense, that is non-linguistic entities. Finally, sometimes he also speaks of changes of *individual people* (a formulation like "not … *only* … single individual[s]" obviously includes them).

In more recent publications, Hacking has distanced himself at least from some aspects of his previous views (see, e.g. Hacking 2006, 2007a,b). While retaining the idea that looping effects are a feature of human and social scientific classifications, he now rejects the concept of an interactive kind. Remarkably, however, the reason he cites for this has nothing to do with the problems just sketched, but rather with his abandonment of the notion of a natural kind. Hacking's thinking about interactive kinds has always been shaped by a commitment to the idea that interactive kinds make sense only as counterparts to natural kinds. As long as he believed that natural kinds constitute a meaningful philosophical concept, this commitment prevented him from extending theories of natural kinds to interactive kinds as well. And when he abandoned natural

kinds from his philosophical terminology, this same commitment led him to abandon interactive kinds.

In my view, neither constructing interactive kinds as a conceptual counterpart of natural kinds, nor excluding natural kinds – or "real kinds", as I prefer to call them – from our philosophical terminology are good ideas. As I shall argue in this paper, interactive kinds are best analyzed as a special case of causal property clusters (CPCs). CPC theory provides an accurate metaphysical conception of real kinds that allows us to conceive of them in a realist manner without succumbing to the pitfalls of essentialism as well as a suitable framework for integrating interactive kinds. Initially, Hacking had "warmly welcomed" CPC theory as having considerable advantages over earlier attempts within his "tradition of natural kinds" (Hacking 1991a,b). However, his commitment to the complementary nature of real kinds and interactive kinds prevented him from applying CPC theory to the latter.

This paper is based on the observation that the discussion about interactive kinds suffers from the fact that neither Hacking nor his critics have worked out a tenable ontology of interactive kinds. The purpose of the paper is to provide just such an ontology. I shall argue that, in spite of Hacking's own doubts, it is possible to maintain the notion of an interactive kind and to develop it in a way that withstands the criticisms raised by other authors. The way to achieve this, however, is to depart from Hacking's "dynamic nominalism" (Hacking 2002, p. 2) and to draw a decidedly realist picture of interactive kinds.

My attempt to reconstruct interactive kinds as CPCs is not without predecessors (see Griffiths 1997, who, however, has some reservations about this project as well as Murphy 2006; Kuorikoski and Pöyhönen 2012, who are broadly sympathetic to it). Unlike these attempts, I shall focus on several basic metaphysical aspects that have not yet been sufficiently addressed. In particular, I shall (in Section 2) provide a detailed examination of the distinction between classifications (which are linguistic entities) and various kinds of kinds (which are non-linguistic entities). In Section 3 I then develop a typology of looping effects according to the ontological sort of entities that are affected. The resulting types are i-looping (individuals), c-looping (classificatory categories), and k-looping (kinds). In Section 4, I introduce the CPC conception of kinds in more detail and argue that it provides a viable ontological framework for modeling interactive kinds. More precisely, I model interactive kinds as a special case of changing CPCs – namely of those CPCs that are subjected to k-looping. In order to do this, a crucial step is to make sense of the more general notion of a changing kind. Finally, in Section 5 I demonstrate the advantages of my proposal over Hacking's own "semantic resolution".

## 2 Real Kinds, Human Kinds, Interactive Kinds, and Classifications

Compared with philosophers of science, Hacking's ideas have not yet received much attention among social ontologists (one notable exception is Mallon 2003). A reason for this relative neglect might be that social ontologists, when dealing with pluralities of human beings, have largely been preoccupied with groups exhibiting collective intentionality or other forms of collectives. Collectives are constituted by intentional or non-intentional relations between their members (mutual acquaintance, regular interactions, etc.); and those relations unite them in a way that is somehow similar to the way material objects are united by relations between their proper parts. By contrast, a *kind* (be it a kind of people – that is, a "human kind" – or any other kind) is constituted not by relations that unite its members (or, more accurately, its "instances") as some sort of object-like whole, but rather by relations of similarity. For example, the set of people exhibiting a certain disease is united by their having encountered the same pathogen and sharing similar symptoms, and not, or at least not necessarily, by any social bonds between them[1] (for a comprehensive study of all possible types of pluralities of human beings, see Hauswald 2014). Nevertheless, there are important respects in which collectives and kinds of people are alike. For example, it is a feature of collectives as well as of kinds of people that their members typically can say "we" when referring to the entity they are, or believe to be, part of.[2] Moreover, there is considerable affinity between Hacking's ideas about human kinds and what social ontologists take to be a fundamental feature of human beings, groups, and the social world in general: what people think about themselves and their social

**1** Of course, all the people belonging to this set might come into contact with one another and establish such social bonds, for example by creating a large self-help support group. Indeed, as several authors have pointed out, different sorts of social pluralities often tend to emerge from one another. Tilly (1978) introduced the term "catnet" (from "category" and "network") to indicate that social networks tend to be comprised of people who exhibit certain commonalities. In a similar vein, Young (1994, p. 735) observed "constant ebbs and flows of groupings out of series". As an example, Young cited feminist groups that are constituted by collective intentionality and that recruit their personnel from the larger plurality of women, which she analyzed as a Sartrean series. Nevertheless, these pluralities remain distinct entities that have a different ontological status. While the self-help support group in my example is *constituted* by the presence of social bonds, the kind (or the set of its instances) is *constituted* by the presence of a variety of commonalities.

**2** Social ontologists usually distinguish between a distributive and a collective reading of "we" (see, e.g. Ludwig 2014). It seems that when speakers refer to the kinds they are instances of, their use of the term "we" always has a distributive sense.

environment can have significant influences on, or be constitutive of, what they are and how social facts are generated.

In philosophy of science, the question of how to conceive of "natural" or "real kinds"[3] plays a crucial role within the debates about realism, laws, and inductive generalizations in science. I adopt a causal property cluster (CPC) view about real kinds, which has proven to be a promising middle ground between essentialism and eliminativism that avoids the pitfalls of both these extreme positions. In essentialism, real kinds are conceived of as ahistorical and eternal, as grounds for exceptionless generalizations. There is mounting evidence that this picture is so strict that it cannot retain the very concept of a real kind. For example, the discussion about ceteris paribus laws has revealed the dubiousness of the idea of exceptionless generalizations even with regard to the kinds studied in the natural sciences (the locus classicus of this debate is Cartwright 1983). But the eliminativists' conclusion that there are no real kinds whatsoever is also problematic because there really are stable (even if not exceptionless) generalizations in all empirical sciences (including the social and human sciences), and some sort of a real kind theory is the most promising strategy to explain this fact. The point is that our conception of kinds must conform to the causal structure of the world and the methodological situation of the sciences. We need a concept that is more "liberal" than the essentialist concept, but still strong enough to account for the inductive success of the sciences. I take CPC accounts to be the best currently available alternatives to meet these conditions. The basic idea behind the CPC view is that individuals belonging to a real kind share a great many properties and that these commonalities are not a matter of pure definition but come about due to causal mechanisms (I will sketch my version of the CPC conception in some more detail in Section 4).

For any form of realism about kinds (and the CPC view is realist in the sense that is relevant here), it is crucial to distinguish between *kinds*, on the one hand, and *classifications* and *classificatory categories*, on the other. One of the main sources of confusion in Hacking's treatment of interactive kinds is that he does not sufficiently acknowledge this distinction. Classificatory categories, as I shall use the term, are linguistic entities. Classifying is a subject's operation of sorting things into classificatory categories or classes and labeling them by certain classificatory terms. By contrast, a kind is a non-linguistic entity. To illustrate this point, if a new disease is discovered, medical scientists will presumably introduce a concept referring to the disease and supplement existing classification systems

**3** I prefer the term "real kind" over the term "natural kind" in order to avoid the impression that the kinds studied in the natural sciences are the only plausible candidates for being real kinds.

such as DSM and ICD with a newly created classificatory category. Obviously, creating a new classificatory category is not tantamount to creating a new kind. For example, tuberculosis (the kind, the disease as such) existed long before it was identified by Robert Koch in the late 19th century as the disease caused by the tubercle bacillus – and, pace Latour (2000), Ramses II might well have died of it, even though an appropriate medical category did not exist in antiquity. Sure enough, classificatory practices may feed back into their targets, or, in extreme forms, even bring them into existence in the first place. After all, these are essentially the phenomena at issue in this paper. However, for one thing, in many "ordinary" cases of classifications there are no such feedback effects. A successful classification is usually thought of as a process that somehow adequately captures some relevant preexisting differences. Moreover, and more importantly, even in cases in which the classification target is somehow produced by the classification process itself, the distinction between the linguistic and the non-linguistic dimension remains crucial. Such cases are best described as processes in which a kind comes into existence as a result of a classificatory practice; and to describe it in these terms is possible only as long as the difference between classifications as linguistic entities and kinds as non-linguistic ones is acknowledged.

Note that the existence of classificatory categories is logically independent of the existence of kinds: some classificatory categories correspond to existing kinds (for example, the category "tuberculosis" corresponds to the disease that is caused by the tubercle bacillus); second, there are kinds without corresponding classificatory categories [e.g. not (yet) known diseases]; finally, there are categories for which there is no corresponding kind. Examples of the latter are provided, on the one hand, by terms like "phlogiston" that fail to capture any existing kind. On the other hand, there may also be categories that do not even purport to single out a kind in any meaningful sense. For example, one could define C as a category that includes exactly the three individuals Shirley Williams, W.V.O. Quine, and Brigitte Bardot.[4] C would be a perfectly intelligible – although probably not particularly useful – classificatory category. But, I suppose, there is no preexisting kind in any realist sense with instances of exactly these three individuals and that could be captured by C. There is not really anything these individuals have in common, except for the properties they also share with others (such as being human).

Also note that an individual may correctly or incorrectly be identified as falling into a category C. For example, a person may be diagnosed as having

**4** Gilbert (1989, p. 149f.) uses this example to illustrate that an arbitrary set of people is not yet a group in the proper sense. Similarly, it is not yet a human kind. To turn a set of people into a group or a kind requires, as argued above, the presence of suitable social bonds or sufficient similarities between them.

tuberculosis, even though she did not encounter the tubercle bacillus – or, in general, even though she does not meet the defining criteria for C. Conversely, someone may not be classified as falling into C, even though she does meet the criteria. These two cases correspond to what physicians and statisticians call false positives and false negatives.

To be brief, I shall use the terms "real kind", "classificatory category", "human kind", and "interactive kind" as follows: A *real kind* is a structure that can be identified with a particular CPC. As a non-linguistic entity, it should not be confused with the *classificatory category* that might be introduced to capture it.[5] A *human kind* is a real kind whose instances are human beings.[6] Hacking sometimes tends to use the terms "human kind" and "interactive kind" interchangeably. I will not mimic this practice. As I use the term, an *interactive kind* is a real kind that is subjected to *kind-looping*, as will be explained in the next section. There is nothing in this definition of an interactive kind, nor in that of a human kind, that suggests that both terms are necessarily co-extensional. In fact, one can easily imagine non-human interactive kinds [e.g. kinds whose instances are (hypothetical) self-aware non-human beings], as well as non-interactive human kinds. The latter may include cases in which the people who are the instances of a human kind K are unaware of this; or cases in which

**5** If a classificatory category "captures" or "corresponds to" a kind, it contains all and only the instances of this kind. Perhaps there are further necessary conditions for a category to successfully capture a kind, but we need not here delve deeper into the question of what these are.

**6** Like Hacking and others, I take human diseases such as tuberculosis to be suitable examples of human kinds. However, even if one accepts that there is a causal property cluster corresponding to tuberculosis, one might object that, strictly speaking, it is not *human beings*, but particular occurrences of the disease kind that are the instances of this kind in the first place, where a particular occurrence of a disease kind D is a person P's particular *state* of being ill, that is, her *having* the disease D. Since I have defined human kinds as real kinds whose instances are *human beings*, this is a legitimate objection. But we can easily refine our definition of a human kind to accommodate it in the following way. A particular occurrence of a disease D is a *dependent entity* – it always exists *in* an individual I (be it a human being or another organism). Whenever there is an occurrence of D, necessarily, there is an individual in which it occurs. We may, then, introduce a distinction between the *primary* and *secondary* instances of a kind. For example, primary instances of the kind *tuberculosis* are particular occurrences of tuberculosis; primary instances of the kind *human being* are individual human beings. In contrast, the notion of a *secondary instance* can be defined as follows: X is a secondary instance of K iff (1) there is a primary instance Y of K and (2) it is metaphysically necessary that Y cannot exist independently of X. While the particular occurrences of a disease D are the primary instances of D, the diseased individuals are its secondary instances. Using this distinction, we may refine our definition of a human kind in the following way: a human kind is a real kind whose primary or secondary instances are human beings (for more details, see Hauswald 2014, Ch. 3.3).

they are aware, but do not react in any relevant way; or cases in which they do react, but their reactions do not provoke any kind-looping, that is relevant changes in K.[7]

As indicated, what distinguishes interactive from non-interactive human kinds, in my view, is the presence or absence of what I call "kind-looping". In order to get a clearer idea of what this means, in the next section I shall provide a typology of looping effects and delineate kind-looping from two other possible forms: individual-looping and category-looping.

## 3 A Typology of Looping Effects

Looping effects, as Hacking understands them, have a relational structure: something interacts with something else. What are the possible relata of the looping relation? One relatum in all cases is relatively uncontroversial: it is the event of one or more person P's believing that they fall into a classificatory category C.[8] Let us call this event E. Note, first, that P's belief may be correct or incorrect (she may or may not really fall into C); and, second, that C may or may not correspond to any real kind K. Obviously then, a looping effect is the changing of something as a result of E. But which entities change? To which ontological sorts do they belong? What is the other relatum of the looping relation?

As far as what the changing entity can be, I shall argue that we should take into account at least three possibilities: (1) the individual person P, (2) the classificatory category C, and (3) the kind K that corresponds to C (if there is such a kind). All of these possibilities can be realized and should be taken seriously. Let us examine the resulting types of looping in some more detail:

1. The first possibility is that the *individual person P* (her behavior, her intentional attitudes, her self-concept; in short: her properties) is (are) changed

**7** For these reasons, I do not think that Hacking (1999, p. 108) was right in claiming that the social/human sciences can be demarcated from the natural sciences on the basis of the distinction between interactive and indifferent kinds.

**8** In some passages, Hacking admits that there are cases in which the conscious reactions of people other than P suffice for there to be a looping effect. For the sake of simplicity, in this paper I am concentrating on looping effects resulting from the self-conscious reactions of the classified people. *Mutatis mutandis*, the results can easily be extended to the other type of looping. For a more comprehensive treatment of the mechanisms that may be at work in cases where the looping is not primarily triggered by self-aware reactions of the classified people, see Kuorikoski and Pöyhönen (2012) from a CPC theory-friendly perspective, and Tekin (2014) from a nominalist viewpoint.

as a result of E. Let us call this form "individual-looping" or "i-looping" for short. It is possible that not only one individual person, but many people $P_1$... $P_n$ falling under C (or believing they fall under C) react in a similar way. In that case, let us speak of "*aggregated* i-looping".

In some cases, i-looping affects the property or properties that are *defining for C*. An example is a (hitherto non-criminal) person who is treated as a "criminal" and, as labeling theory predicts, changes her self-concept or her behavior in a way that conforms to the role-model of a criminal and ultimately really commits crimes. This is the self-fulfiling-prophecy variant of this case. There is also a self-destructive variant. As an example, consider someone who becomes aware that she suffers from a certain disease (that is, falls into a certain disease category C) and, as a consequence, attempts to overcome it and become cured. If she succeeds, she no longer falls into category C.

In both the self-fulfiling and the self-destructive variant, the looping affects the properties that are defining for C (the property of being criminal or having the disease), which means that the looping has the result that the person either begins or ceases to fall into C. There are also "milder" forms of i-looping in which other properties of P are affected, but not the properties that are defining for C. For example, a person who has reached a certain age and becomes classified as "an adult" might, as a result, change some of her attitudes (she "feels grown-up" now, is proud not to be a child anymore, behaves in a different way, etc.). In this example, the person's membership in the category ("adult") is not affected by her self-aware reactions; she now simply belongs to it due to her age, regardless of her reactions. What her reactions affect are a number of her other properties (her attitudes, behaviors, etc.).

2. In some cases, a change in the category C can be an outcome of E. Let us call these forms "category-looping" or "c-looping". As opposed to the examples just discussed, in cases of c-looping it is not (or not only) the properties of individual people that are affected by the looping. Rather, c-looping affects how a classificatory category is actually conceived of or defined in the relevant community.

   For example, consider homosexuality. Up until a few decades ago, homosexuality was defined as a mental illness. The fact that it was later removed from almost all accepted catalogs of mental disorders can, at least partly, be explained as a result of looping effects. People who fell into the category of a homosexual sought to convince society that it is not pathological and should not be defined as a disorder. The fact that they ultimately succeeded turns this case into an example of c-looping.

3. Finally, it seems that the *kind* that corresponds to the category C can itself change (an extreme case is when the kind eventually ceases to exist). This presupposes, of course, that such a kind actually exists; for, as I made clear in the previous section, not every classificatory category captures an existing kind. However, it is also conceivable that it is not a preexisting kind which changes, but instead that a hitherto non-existing kind simply comes into existence due to looping effects. For example, a number of people might collectively believe that they fall into a classificatory category C which they falsely believe to correspond to a certain kind K. Due to aggregated i-looping, these people might develop a certain characteristic pattern of behavior, emotions, attitudes, etc., so that ultimately a completely new kind of people might be said to have evolved. At this point, I do not want to commit myself to the view that such processes really occur, and I will not further elaborate on them in this paper. But we should at least take this scenario into account as another conceptual possibility. Let us call the changes (or the genesis) of kinds that arise due to the reactions of those who correctly (or, if there is not yet any corresponding kind, falsely) believe themselves to be instances of them "kind-looping" or "k-looping".[9] Kinds that are subjected to k-looping would then be "interactive kinds" in a strict sense.

   What about the relationship between k-looping and aggregated i-looping, and more generally, between changes of a kind and the aggregated changes of its instances? Traditionalists may hold that while a kind's instances can surely change, kinds as such are incapable of genuine change. Others may be inclined to think that a change of a kind is nothing more than an aggregated change of its instances. In the next section, I shall argue that, contrary to traditional metaphysical thinking about kinds, the CPC view provides a framework in which changing kinds can be modeled in a rather straightforward and natural way. But there are good reasons to suggest that aggregated changes of some kind's instances are neither necessary nor sufficient to change the kind (in brief, the evolution of species indicates that it is not necessary; the possibility of aggregated, but merely superficial changes of a kind's instances suggests that it is not sufficient). In any case, we need to enhance our awareness of the variety of ways in which kinds and their instances can change. Having explored how changing kinds can be modeled within the CPC framework, I then go on to interpret interactive kinds as a special case of changing CPC kinds – namely, those kinds that change as a result of reactions of the people who come to believe that they are instances of them.

**9** "Believing to be an instance of K" is equivalent to "believing to belong to a classificatory category that captures K".

# 4 Interactive Kinds as Changeable Causal Property Clusters

If there are any true interactive kinds, they obviously form a subclass of changing kinds. In the following, I shall examine, first, what changing kinds are; and second, what it means for a changing kind to be interactive rather than non-interactive.

Let us first determine the sorts of changes that a kind may be subjected to. First, there are *quantitative* changes in the sense that a kind may gain or lose instances. This is a sort of change that even many traditional prime examples of natural kinds, such as chemical elements or elementary particles, are subjected to. For example, the amount of any particular chemical element in the universe is constantly changing due to processes such as nucleosynthesis and nuclear decay. Second, there are *qualitative* changes: at a certain point in time, the instances of a kind may be qualitatively different from those at another point in time. Qualitative changes can come in different strengths. Suppose, on the one hand, that scientists mark all existing specimens of a certain biological species with pink dots. On the other hand, consider the case of evolutionary developments. Suppose that the same species adapts to new environmental challenges in such a way that the later individuals have pink dots (just like those from the previous example) where their ancestors had none. What distinguishes theses two examples? One difference is that, while in both cases the individuals differ from the individuals of earlier generations, the individuals in the second scenario did not (or not necessarily) change themselves. Evolution is brought about by mutation and selection, not by Lamarckian adaptations. This is why changes of a kind are conceivable without aggregated changes of its instances. A second difference is that, in a way, the evolutionary changes are more pervasive and stable – not in the sense that they are fixed once and for all (of course, evolution goes on and new environmental conditions may provoke further adaptations), but in the sense that the changes are brought about by changes in the regular mechanisms that are responsible for many of the individuals' traits. If the pink dots are merely painted, that is, if they are a non-heritable trait, we would expect that when a new specimen is born it will be "normal" in the sense that it does not have a pink dot. If the pink dots are due to evolutionary changes, that is, to a heritable trait, they will be passed to the offspring. Also, in the evolutionary scenario, the number of properties that differ between individuals before and after the adaptation will presumably be higher than when the pink dots are merely painted. If an individual is marked by a painted dot, it changes in just this property of having or not having the dot. But if the pink dots are due to evolutionary adaptations, there will be a number

of further differences between the pre- and the post-adaptation individuals, most notably differences in the genotype.

These examples suggest that a kind can undergo weaker and stronger sorts of changes. One might be inclined to say that only some of the examples – perhaps only the last one involving the evolutionary scenario – represent true or genuine changes of the kinds as such. And to account for this case, one could introduce a distinction between the genuine and non-genuine changes of a kind, where the former might include qualitative changes similar to those of the evolutionary scenario and the latter quantitative changes and qualitative changes similar to those of the painted-dot scenario. However, there may also be contradictory intuitions,[10] and, arguably, it will be difficult to establish general criteria for determining when something is a genuine rather than a non-genuine change. For the purposes of this paper, it suffices to make a gradual distinction between weaker and stronger forms of changes of a kind, and whoever wants to demarcate genuine from non-genuine changes may define a threshold on this continuum wherever they believe it to be appropriate. Possible criteria for determining whether something is a weaker or stronger qualitative change could, arguably, include: the number of properties changed, the extent to which they are changed (consider for example a property such as the specimens' size, which may change to a greater or lesser degree), and the stability and pervasiveness of the change (is it brought about by fundamental changes in the regular mechanisms underlying the traits of the kind's instances – as in the evolutionary example – or by more superficial influences – as in the painted-dot example?). Concerning quantitative changes, weaker and stronger forms can obviously be distinguished by taking into account the number of instances that a kind loses or gains.

How can we make metaphysical sense of a kind undergoing these various forms of change? In the following I will introduce the CPC view as a framework in which changing kinds can be modeled in a rather straightforward and natural way. Let me first briefly sketch the basic ideas behind the CPC conception of kinds. According to the CPC view, "real" (or "natural" or "scientifically significant") kinds correspond to CPCs. The core idea dates back to 19th century philosophers such as William Whewell and John Stuart Mill. According to Whewell, individuals belonging to a real kind are characterized by "an inexhaustible body of resemblances amongst individuals … made by nature, not by mere definition"

**10** For example, as concerns quantitative changes, it might seem natural to hold that a kind does not undergo a genuine change just because it has more or less instances than it had before. On the other hand, one might argue that every kind has a property P of being instantiated so-and-so many times. Provided that an entity X changes whenever there is a change in its properties, a kind *does* change when it gains or loses instances, since it changes at least in relation to P.

(Whewell 1860/1971, p. 290). Mill writes that "a hundred generations have not exhausted the common properties of animals or of plants, of sulfur or of phosphorus; nor do we suppose them to be exhaustible, but proceed to new observations and experiments, in the full confidence of discovering new properties which were by no means implied in those we previously knew" (Mill 1843/1973, p. 122). In these passages, Whewell and Mill express two ideas that are crucial to the conception of a real kind: the idea that the instances of such a kind have a large number of properties in common and the idea that the similarity between them is a result of causal processes, not of "mere definition".

The CPC conception of kinds captures these ideas.[11] Here is the basic model. Imagine a multidimensional space of properties (MSP) in which every dimension represents a property and in which all existing individuals are located. The nearer two individuals are in this MSP, the more properties they share, that is, the more similar they are. In the MSP, individuals are neither homogeneously nor randomly distributed. Instead, their distribution is structured. In some areas there are many individuals; other areas are empty. I call the latter areas "realization gaps" (this expression is borrowed from Pinkal 1995, p. 106), the former ones "realization accumulations". The accumulations can be identified with real kinds. Moreover, the co-occurrence of properties in the individuals belonging to the kinds is not due to mere definition, but due to causal mechanisms. Consider biological species as an example. The individuals that instantiate a particular species differ in many ways – no cat perfectly resembles any other cat. Thus, no two cats occupy the very same location in the MSP. At the same time, due to causal mechanisms such as heredity, all cats share a great many properties with many other cats. This means that even though no two cats occupy perfectly identical locations, there is a relatively small region in the MSP where all cats are located. Other species occupy other regions in the MSP. In between these regions there are realization gaps, which are not occupied by any individuals.

Two things need further emphasis. First, since more or less properties can cluster together and different realization accumulations can be more or less clearly separated from each other by realization gaps, the notion of a real kind turns out to be gradable. Some kinds are "more real" (or "more natural", more "scientifically significant") than others. Second, the CPC conception is not

**11** Apart from Mill and Whewell, other members of Hacking's (1991a) "tradition of natural kinds", such as Venn, Peirce, Russell, and Quine, can also be regarded as predecessors of the CPC approach. For current versions, see, e.g. Boyd 1991, 1999, 2010; Kornblith 1993; and Hawley and Bird 2011. I am not relying here on any particular CPC theory, but use a version that is rather neutral (without, however, committing myself to the claim that it is compatible with all existing CPC theories).

committed to essentialism in the sense that it is not necessary to assume that membership in a kind is possible only when an individual has one or more putative essential properties. What is required, however, is that a kind exhibits a sufficient number of the properties of the cluster. But no single property needs to be considered as necessarily occurring in each individual that instantiates the kind.

We are now in a position to answer the question of what it means for a kind to change, and what it means for a changing kind to be interactive. Within the framework of the CPC conception, *qualitative* changes of a kind can be modeled simply as movements of the realization accumulations through the MSP. The more extensively the properties change, the further the accumulation moves. For example, in the case of evolving biological species, the properties of the individuals that instantiate a species change over time. Individuals of the species at an earlier point in time exhibit a certain cluster of properties, whereas the individuals of the species at a later point in time exhibit partly different properties: the realization accumulation has moved from one region in the MSP to another.[12] In the case of mere *quantitative* changes, the accumulation does not move, but instead fluctuates in size or "density".

*Kind-looping* can now be reconstructed as a special case of changes of a kind: namely, as a case in which a quantitative or qualitative change comes about as the effect of individuals who realize or come to believe that they are instances of the kind. *Interactive kinds* are kinds that are subjected to these changes. It is reasonable to assume that the notions of an interactive kind and of kind-looping inherit the gradualness that we have found to be characteristic of the notion of a change of a kind. If quantitative changes are only very weak forms of change (or no genuine change at all), loopings that merely involve quantitative changes will be only very weak forms (or no genuine cases) of kind-looping as well. As an example of this quantitative case, consider a hitherto incurable human disease kind that loses instances because many people who suffer from the disease (and who know that they do) learn about a new successful therapy, try it out, and are eventually cured.

For an example of qualitative kind-looping, consider so called culture-bound syndromes, that is, medical conditions that are restricted to particular cultural

**12** This raises difficult questions about how to individuate CPCs/kinds: when a cluster moves through the MSP, up to which point does it remain numerically the same? Although I shall not delve deeper into this problem, let me note that the CPC conception itself does not suggest a general answer. One will have to relate a different story about biological species than about chemical elements. But in both cases, arguably, there remains an element of convention that cannot be eliminated. As well as being made by nature, kinds are also the "workmanship of women and men" (Boyd 1999).

contexts and believed to be caused by the socio-cultural particularities of these contexts (American Psychiatric Association 2000, p. 897ff.). If the socio-cultural conditions that are responsible for shaping the syndrome are altered in such a way that the symptoms are changed, we obviously have a rather strong example of a change of a kind – somewhat on a par with the evolutionary example discussed above. And if the alterations in the relevant socio-cultural conditions occur as a result of the self-aware reactions of those who have the disease, we would have a strong example of kind-looping.

In the next and last section, I shall further clarify my approach by comparing it with the position Hacking takes in his *The Social Construction of What?*

## 5 Some Remarks on Hacking's "Semantic Resolution"

In this section, I will criticize the "semantic resolution" Hacking proposes in response to the intuition that some kinds seem to be affected by looping in some respects, but not in others; or, as Hacking himself puts it, they seem to be interactive and indifferent kinds at the same time (see Hacking 1999, p. 108ff.). I argue that Hacking's solution is unconvincing as he does not adequately differentiate between different types of looping effects and implicitly adopts an inadequate ontology of kinds.[13]

Hacking's example is autism. He assumes that it might well be that the kind autism "*is* (is identical to) a certain biological pathology P" (Hacking 1999, p. 119; emphasis in original) and then asks himself how it can still be considered an interactive kind. The answer he gives draws on Putnam's semantics of natural kind terms, in particular his notion of a stereotype. Hacking suggests that while the stereotype associated with autism in previous decades differs from that associated with it nowadays (because new medical knowledge has been created, social discourses about autism have changed, etc.), the underlying structure P remains the same. Drawing on this assumption, Hacking now proposes that what makes autism an interactive kind is the fact that the stereotype associated with it has changed. This, however, is unconvincing. The stereotypes associated with almost all kinds will change over time as a result of new scientific discoveries and changing common assumptions about these kinds. This holds true for diseases and other human kinds as well as for gold, electrons, or

**13** This criticism is akin to Murphy's (2006, p. 271ff.), who, however, neither explicitly grounds it in a CPC conception nor relates it to a typology of looping effects, as I shall do.

planets. Stereotypes are just part of our cognitive representations of the kinds in question. But of course, if our stereotype of, say, gold changes, this does not imply that the kind gold itself changes. If the stereotype associated with autism changes as a result of how autistic people reacted to the way they are treated and what is believed about them in society, this alone would merely amount to a case of *category-looping*, as defined in Section 3; that is, a case of looping that affects how a classificatory category is conceived of or defined in the relevant community. To claim that autism is an interactive *kind* would require demonstrating the presence of kind-looping.

However, by drawing on Putnam's model of natural kind terms, Hacking implicitly imports his essentialism. This becomes apparent in his identification of autism with an underlying pathology P. The problem is that essentialism does not really provide the conceptual resources needed to model how autism – or any other kind – might be subjected to strong forms of change, and a fortiori of kind-looping. Essentialists can make sense of quantitative changes and weak qualitative changes, but not of stronger forms of change. Essentialists would have to say that a strong form of change in a kind requires a change of its essence (for the kind *is*, is identical to, its essence). But an essence is, by its very nature, a *necessary* feature, and something that fails to have the essence fails to be an instance of *that* kind. Individuals at a later point in time either still have the same essence as individuals at earlier points in time or they do not have the same essence. In neither case it is possible to speak of a changing kind in the strong sense; for, in the first case, the kind (which is identical to the essence) did not change, and in the second, the former and the later individuals are just of completely different kinds.

If we abandon essentialism and adopt a CPC view, as I have suggested in the preceding sections, we can still be realists about kinds and at the same time be able to conceive of strong forms of change as well as of kind-looping. A qualitative change of a kind like autism can then be reconstructed as a movement of the realization accumulation corresponding to this kind through the MSP. On the CPC view, no essences have to be postulated, and the property cluster can easily be conceived of as changeable. The cluster can undergo qualitative changes and still remain numerically the same.

Autism can be shown to be an example of an interactive kind in the strong sense if it can be shown that (1) autistic people, at later points in time, have different properties than they (that is, the very same people or other autistic people) had at earlier times, (2) particularly many properties change and/or this change of properties is due to changes of the mechanisms that are responsible for the autistic CPC, and (3) these changes are at least partially provoked by self-aware reactions of the autistic people themselves. I leave open whether these conditions

really are fulfiled, since this would require answering empirical questions that go beyond the scope of this article. What is important, however, is that the CPC conception provides a sound conceptual framework for modeling what goes on in interactive kinds. Autism is subjected to strong kind-looping if the conditions 1–3 are met. If only the conditions 1 and 3 are met, but not condition 2, perhaps we would still have an example of kind-looping, but only a weak form of it. If the stereotype associated with autism changes, or more generally, if the categories and concepts with which autism is conceptualized are modified (as has happened, for example, when a number of diagnostic categories including "autistic disorder", "Asperger's disorder", "childhood disintegrative disorder", "Rett's disorder", and "pervasive developmental disorder not otherwise specified", which were differentiated in DSM-IV, were conjoined into a single category labeled "Autism Spectrum Disorder" in DSM-5; see American Psychiatric Association 2013, p. 809), and if these modifications are caused or influenced by the reactions of autistic persons, we would have a case of c-looping.

## 6 Conclusion

I have argued that it is possible to retain a meaningful notion of an interactive kind in spite of the criticisms that have been raised against it. However, in order to do so, major departures from Hacking's original approach are necessary. I have opted for a realist interpretation of interactive kinds, according to which interactive kinds are not the opposite, but instead comprise a subclass of real kinds. If we conceive of real kinds as CPCs (and we should because the CPC view is superior to alternative theories of kinds such as essentialism and eliminativism), it is possible to model, first, changes of kinds as movements of realization accumulations through the MSP. Second, we can model interactive kinds as those changing kinds in which these movements are due to self-aware reactions of the people who instantiate the kinds.

It was not the primary aim of this paper to argue that looping effects are a ubiquitous and unavoidable feature of the social and human sciences or that all the kinds studied by them are interactive kinds. Although I do suspect that all types of every different strength of looping effects occur more or less frequently – often overlapping because people, classificatory categories, and kinds are likely to change in parallel – a valid proof of this conjecture would require comprehensive empirical considerations that go beyond the scope of this paper. Instead, I hope to have provided a sound conceptual and ontological framework to model such effects, provided they exist.

# Bibliography

American Psychiatric Association (ed.) (2000): *Diagnostic and Statistical Manual of Mental Disorders. 4th Edition Revised. DSM-IV-TR.* Washington, DC: American Psychiatric Association.

American Psychiatric Association (2013): *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition. DSM-5.* Washington, DC: American Psychiatric Association.

Becker, H. S. (1963): *Outsiders: Studies in the Sociology of Deviance*. New York: The Free Press.

Berreby, D. (2005): *Us and Them: Understanding Your Tribal Mind*. New York: Little, Brown.

Boyd, R. (1991): "Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds". In: *Philosophical Studies* 61. No. 1/2, p. 127–148.

Boyd, R. (1999): "Kinds as the 'Workmanship of Men': Realism, Constructivism, and Natural Kinds". In: J. Nida-Rümelin (Ed.): *Rationality, Realism, Revision*. Berlin: De Gruyter, p. 52–89.

Boyd, R. (2010): "Realism, Natural Kinds, and Philosophical Methods". In: H. Beebee and N. Sabbarton-Leary (Eds.): *The Semantics and Metaphysics of Natural Kinds*. New York: Routledge, p. 212–234.

Cartwright, N. (1983): *How the Laws of Physics Lie*. Oxford: Clarendon Press.

Cooper, R. (2004): "Why Hacking is Wrong about Human Kinds". In: *British Journal of Philosophy Science* 55. No. 1, p. 73–85.

Dupré, J. (2002): *Humans and Other Animals*. Oxford: Clarendon Press.

Gilbert, M. (1989): *On Social Facts*. London: Routledge.

Griffiths, P. E. (1997): *What Emotions Really Are*. Chicago: University of Chicago Press.

Hacking, I. (1991a): "A Tradition of Natural Kinds". In: *Philosophical Studies* 61. No. 1/2, p. 109–126.

Hacking, I. (1991b): "On Boyd". In: *Philosophical Studies* 61. No. 1/2, p. 149–154.

Hacking, I. (1995): "The Looping Effect of Human Kinds". In: D. Sperber, D. Premack, and A.J. Premack (Eds.): *Causal Cognition: A Multidisciplinary Debate*. Oxford: Clarendon Press, p. 351–383.

Hacking, I. (1998): *Mad Travellers, Reflections on the Reality of Transient Illnesses*. Charlottesville and London: University Press of Virginia.

Hacking, I. (1999): *The Social Construction of What?* Cambridge, Mass: Harvard University Press.

Hacking, I. (2002): *Historical Ontology*. Cambridge, Mass: Harvard University Press.

Hacking, I. (2006): "Types De Gens, Des Cibles Mouvantes". http://www.college-de-france.fr/media/ian-hacking/UPL4141595277092170127_6___Types_de_gens_des_cibles_mouvantes.pdf, visited on 16 October 2015.

Hacking, I. (2007a): "Kinds of People: Moving Targets". In: *Proceedings of the British Academy* 151. p. 285–318.

Hacking, I. (2007b): "Natural Kinds: Rosy Dawn, Scholastic Twilight". In: *Royal Institute of Philosophy Supplement* 61. p. 203–239.

Hauswald, R. (2014): *Soziale Pluralitäten. Zur Ontologie, Wissenschaftstheorie und Semantik des Klassifizierens und Gruppierens von Menschen in Gesellschaft und Humanwissenschaft.* Münster: mentis.

Hawley, K and Bird, A. (2011): "What are Natural Kinds?". In: *Philosophical Perspectives* 25. No. 1, p. 205–221.

Hirschfeld, L. A. (1996): *Race in the making: Cognition, Culture, and the Child's Construction of Human Kinds*. Cambridge, Mass.: MIT Press.

Khalidi, M. (2010): "Interactive Kinds", *British Journal of Philosophy Science* 61. No. 2, p. 335–360.
Kornblith, H. (1993): *Inductive Inference and Its Natural Ground*. Cambridge, Mass.: MIT Press.
Kuorikoski, J and Pöyhönen, S. (2012): "Looping Kinds and Social Mechanisms". In: *Sociological Theory* 30. No. 3, p. 187–205.
Latour, B. (2000): "On the Partial Existence of Existing and Non-existing Objects". In: L. Daston (Ed.): *Biographies of Scientific Objects*. Chicago: University of Chicago Press, p. 247–269.
Ludwig, K. (2014): "The Ontology of Collective Action". In: S.R. Chant, F. Hindriks, and G. Preyer (Eds.): *From Individual to Collective Intentionality*. Oxford: Oxford University Press, p. 112–133.
Mallon, R. (2003): "Social Construction, Social Roles, and Stability". In: F. Schmitt (Ed.): *Socializing Metaphysics: The Nature of Social Reality*. Lanham: Rowman & Littlefield, p. 327–354.
McNally, R. J. (2011): *What is Mental Illness?* Cambridge, Mass: Belknap Press of Harvard University Press.
Mill, J. S. (1843/1973): *A System of Logic*. Toronto: University of Toronto Press.
Murphy, D. (2006): *Psychiatry in the Scientific Image*. Cambridge, Mass: MIT Press.
Pinkal, M. (1995): *Logic and Lexicon: The Semantics of the Indefinite*. Dordrecht: Springer.
Tekin, S. (2014): "The Missing Self in Hacking's Looping Effects". In: H. Kincaid and J. Sullivan (Eds.): *Classifying Psychopathology: Mental Kinds and Natural Kinds*. Cambridge, Mass.: MIT Press, p. 227–256.
Tilly, C. (1978): *From Mobilization to Revolution*. New York: McGraw-Hill.
Tsou, J. (2007): "Hacking on the Looping Effects of Psychiatric Classifications: What Is an Interactive and Indifferent Kind?". In: *International Studies in the Philosophy of Science* 21. No. 3, p. 325–340.
Whewell, W. (1860/1971): *On the Philosophy of Discovery*. New York: Franklin.
Williams, A. (2015): *Forensic Criminology*. New York: Routledge.
Young, I. M. (1994): "Gender as Seriality: Thinking about Women as a Social Collective". In: *Signs Journal of Women in Culture and Society* 19. No. 3, p. 713–738.