Social categories in the making: construction or recruitment?

Samuli Reijula

Received: 31 October 2020 / Accepted: 23 July 2021 / Published online: 2 August 2021
© The Author(s) 2021

Abstract
Real kinds, both natural and social categories, are characterized by rich inductive potential. They have relatively stable sets of conceptually independent projectable properties. Somewhat surprisingly, even some purely social categories (e.g., ethnicity, gender, political orientation) show such multiple projectability. The article explores the origin of the inductive richness of social categories and concepts. I argue that existing philosophical accounts provide only a partial explanation, and mechanisms of boundary formation and stabilization must be brought into view for a more comprehensive account of inductively rich social categories.

Keywords Social category · Social kind · Multiple projectability · Social construction

1 Introduction: social categories as substance concepts

This article addresses a puzzle concerning social kinds. Just like natural kinds, several social kinds—from ethnic groups and professions to political orientations—are characterized by rich inductive potential. In other words, they have relatively stable sets of conceptually independent projectable properties. But why do these clusters form, why is there such clotting in the social world? I argue that existing philosophical accounts cannot fully explain the inductive potential of many social kinds. I propose that new a social kind and its “thingness” often result from processes of (strategic) boundary formation, which can be pictured as existing differences being stitched together or coalescing into new distinct categories.

This article belongs to the topical collection “Concept Formation in the Natural and Social Sciences: Empirical and Normative Aspects”, edited by Georg Brun, Catherine Herfeld, and Kevin Reuter.

Samuli Reijula
samuli.reijula@helsinki.fi

1 Theoretical Philosophy, University of Helsinki, PL 24, 00014 Helsinki, Finland
Scientific concepts come in various flavors. They may be, for example, definitions, quantitative, or classificatory (see Hempel, 1964). In this article, I focus on just one central kind of scientific concept and its role in the epistemology of science. Ruth Millikan has called them substance concepts. By substances, she means things that retain their properties over our numerous encounters with them. This makes it possible for us to store away knowledge or knowhow concerning them (obtained via our encounters with their members) for use on later occasions. Furthermore, this epistemic affordance should not be an accident, but should be supported by an ontological ground of real connection (Millikan, 2000, p. 2).

Natural kinds are a central example of substance concepts employed in science. Natural kind concepts serve various roles in research (Pöyhönen, 2014): First, they coordinate research activities: establishing \( K \) as a natural kind makes it a sensible target for scientific research, because assuming natural kindhood gives us a reason to believe that, in addition to the knowledge we already possess about the members of the kind, there are yet more things to know about (precisely) the group of things referred to by the concept. Several recent discussions on natural kinds in the philosophy of science reflect this idea: our scientific concepts should be revised to capture genuine divisions in reality so as to best serve our scientific practices of prediction, explanation and intervention (Craver, 2004; Machery, 2009).

Millikan’s distinction between kinds and properties reflects this coordinative function: what distinguishes a kind from a property is that a kind carries various properties. In a sense, substance concepts act as nodes around which scientific knowledge gets accumulated. What matters is not only what we already know about a particular kind. Hierarchical systems of categories inform us about what sort of things we could know. For example, based on our knowledge of physical stuffs, we know that the spatiotemporal location of a particular instance of the kind ‘gold’ is not one of its invariant properties, whereas its atomic number is. Knowing the kind of substance that gold is, we know the same applies to other elements as well, irrespective of whether we already know their atomic numbers or not.

The fact that kinds (more or less) reliably carry their properties from one encounter to the next is of particular interest for predictive, explanatory, and manipulation-grounding reasoning. Establishing the kind membership of a particular object unlocks several other inferences regarding it. If, for instance, after observing some diagnostic combination of properties, we identify a particular nugget of stuff \( S \) as gold, we have good reasons to believe, for example, that it would not dissolve in nitric acid, and that one gram of \( S \) could be drawn into a metal wire up to 66 km in length. Likewise, identifying an animal as a tiger gives good reasons to expect it has four legs, stripes, and DNA sequences characteristic of tigers. This property is what Godman (2015) calls multiple projectability: members of real kinds have numerous (perhaps innumerable) properties in common, and therefore, they enable a broad range of empirical generalizations.

Such inferences are rarely certain. Typically, they are only probabilistic, as albino tigers and ones with only three legs do exist. There are also differences between kinds, and kinds of kinds, in the number of inferences licensed by the identification of the kind membership. Many substance concepts utilized outside the physical and chemical sciences would not satisfy the criteria of natural kindhood for traditional
essentialist or nomological theories of natural kinds. They can, however, manifest multiple projectability and serve a similar role in scientific inference. In this sense, it is possible that biological, social, and behavioral categories may be called real kinds and the concepts referring to them, substance concepts.

In this article, I focus on a particularly unlikely case of substance concept, namely, social categories. Strong stereotypes and prejudices are often associated with categories such as gender, ethnicities, and professions. As I argue below, despite stereotypic beliefs often being false, it would be a mistake to treat them as conventional classes, in contrast to inductive useful kinds. Just as in the natural sciences, we have good reasons to treat many social categories as substance concepts. Their members are similar in various, often nontrivial ways. Where does such inductive richness originate? I argue that existing philosophical accounts provide only a partial explanation: to give a more comprehensive account of inductively rich social categories, mechanisms of boundary formation and stabilization must be brought into view.

The article proceeds as follows. In Sect. 2, I introduce my example, the puzzling correlation in the US between political orientation and apolitical lifestyle choices, and set up the multiple projectability puzzle. Sections 3–5 discuss existing solutions and refine the formulation of the puzzle. Section 6 shifts the perspective; in it, I introduce an alternative theoretical account of the grounds of multiple projectability of social categories. In Sect. 7, I examine the implications of this account for questions concerning the reality of social categories and their status as real kinds.

2 The puzzle of lifestyle politics

In an article entitled Why do liberals drink lattes? (published 2015 in the American Journal of Sociology) Daniel DellaPosta, Yongren Shi and Michael Macy show that some cultural stereotypes have substantial predictive value: in the US, political liberals tend to drink lattes, drive Volvos, enjoy modern art and reggae music, whereas those with conservative political views report a preference for bird-hunting and less trust in science.\(^1\) Earlier research in cultural sociology had already confirmed various correlations between political ideology and views on normatively loaded topics such as gay rights, gender roles, school prayer, and abortion, but the DellaPosta study, based on broad empirical evidence from the General Social Survey (1972–2010), found that political views are associated with prima facie nonideological lifestyle properties such as the ones mentioned above.

Consequently, ‘liberal’ and ‘conservative’ designate social categories with surprisingly rich inductive potential: Finding out about a person’s political orientation gives rise to various justified expectations concerning their views on moral matters and lifestyle choices. Unsurprisingly, the magnitude of correlations is typically

\(^1\) Curiously, despite holding a more positive view of science, liberals are also more likely to read horoscopes and to believe in the “supernatural powers of deceased ancestors” (DellaPosta et al., 2015, pp. 1482–1483).
not large, but identifying a person’s political orientation does provide some inductive mileage for predicting other things about that individual. In other words, in the United States, social categories associated with political orientation appear to be characterized by the multiple projectability of substance concepts.

How could we explain the multiple projectability of social categories? The first option that may come to mind is a Kantian one: It is us and our conceptual activity that leads us to perceive categorical structure in the social world. In earlier contributions to the tradition of natural kinds, social classifications were sometimes portrayed as *conventional classes*: In contrast to concepts standing for genuine divisions in reality, the unity of conventional classes is conceptual or linguistic in nature. An example of a conventional class could be the one consisting of the Eiffel Tower and the things on the left side of my desk. Although nothing prevents us from using such a concept and even labeling such a strange kind, it is radically different from social categories. Conventionism would make our inductive success using such social categories a miracle. Instead, the most natural explanation for us perceiving similarities between the members of social kinds is that there *are* objective similarities between them. Hence, although the concepts referring to structures in the social world are certainly of our own making, the categories themselves may be objective, or real, in the sense that they and their properties are not up to any single individual to decide. Nor are social categories transparent to us in a sense that conventionalism would suggest: social kinds often have stable properties not obvious to all (or any) member of the linguistic community. Consequently, there are discoveries to be made about them. These considerations rule out the “error theory” approach to social categories, their unity cannot lie solely in the eye of the beholder.

That said, essentialist or nomological theories of natural kinds (see Hacking, 1991; Murphy, 2006, chapter 9), when applied to social categories, lead to an equally incredible picture of the grounds of projectability. The heterogeneous cluster of properties characteristic of liberals or conservatives does not seem to flow from any internal essence, nor does it appear to be held together by anything that could be called laws of nature. Social categories simply do not appear to latch onto a timeless unit or niche in the structure of reality.

But could it be that the cultural profiles associated with political ideology reflect underlying biological differences between the members of the two categories? Such an explanation has been proposed by the moral foundations theory (Haidt, 2012). The theory suggests that whereas the liberal moral palate is sensitive to considerations of care, fairness, and liberty, conservative views also reflect three other values: loyalty, authority and sanctity. If these values are understood as arising from our make-up as biological organisms, the political orientations can be seen as *hybrid* or *biosocial kinds* (Mallon, 2016; Murphy, 2006).

In the case described by DellaPosta and coauthors, an explanation of the lifestyle enclaves that would refer to demographic or biological background variables could, at best, only explain some aspects of the rich cultural profile associated with each political orientation. Especially the associations between political orientation and nonideological lifestyle choices seem difficult to explain by such theories. Instead, the cluster of properties, although real, seems contingent with respect to our make-up as biological organisms. As DellaPosta and coauthors (2015, p. 1502) point out,
the “Woodstock generation” might have shaved their heads, had it been the supporters of the war who had long hair.

More generally, although the unity of some social categories is explained by shared biological factors, it would be far-fetched to claim that this is the case for all social kinds. Therefore, I set such biobehavioral kinds aside and focus on what could be called purely social categories, for which the unity of the kind, and its difference from other (contrastive) categories is not explainable by features best described in terms of biological vocabulary. For the present purposes, the puzzle of lifestyle politics described by DellaPosta et al. illustrates the more general puzzle of the multiple projectability of purely social categories.

3 The collective acceptance approach and the opaqueness problem

In his (1995) book *The Construction of Social Reality*, John Searle put forward a seminal account of the nature of social reality. Many of the institutional phenomena that Searle discusses are coextensive with what we have here called purely social categories, so it seems natural to apply Searle’s approach to the puzzle at hand. In Searle’s analysis, natural things (X) get associated with social things and proper-

i

ties (Y) through status functions. A status function of the form “X counts as Y in context Z” specifies the relationship between a social thing (a status) and a natural thing. Irrespective of who explicitly announces the status function, or whether such explicit announcement is identifiable at all, statuses gain their stability and reality through the collective acceptance of the community. Furthermore, statuses come with deontic powers: They specify rights and obligations that go together with the social roles and institutions in question. For example, if a piece of metal counts as money, we can use it as payment to obtain goods and services. Or a person assigned the social status of being a priest can declare two people as being married. The status of being married, in turn, has consequences for the couple. Their taxes might be calculated differently, for example. On Searle’s account, such hierarchical structures of institutions weave the fabric of social reality.

It appears that by treating the set of deontic powers and properties implied by the status as the kind’s projectable set of properties, we get a candidate explanation for multiple projectability. And surely Searle’s account gets us one step closer to understanding how purely social categories can acquire inductive richness. But the collective acceptance account cannot provide the full picture of social categories such as sex, gender, or being a refugee: “[M]embership in a certain human kind comes with constraints and enablements that are not justified with reference to the presence of the property that is taken to define the kind.” (Asta, 2013, p. 729). For example, refugees face constraints and enablements which are not logical consequences of the legal privileges and burdens that come with the legal status itself.

Not all properties of purely social categories directly result from institution-grounding speech acts or collective acceptance. Uncovering that whole range should not, however, even be seen as the aim of Searle’s account. Rather, his analysis is intended to capture the “logical structure of society.” From our current point of view, that is exactly its weakness, because according to the account, the logic of
institutions is ultimately grounded in the act of collective acceptance. Alone, acceptance, or any kind of mental activity for that matter does not generate social reality. We need to look at actions and practices. As theorists of the social world from Anthony Giddens and Pierre Bourdieu to Barry Barnes have emphasized, social reality must be constantly reproduced and re-enacted by our actions and practices. Often such practices have consequences not intended or even perceived by the practitioners.

Our example of the categories associated with political orientation serves as a fitting illustration of the shortcomings of the collective acceptance approach: Bird-hunting and a preference for Italian-style milk coffee appear to be apolitical properties that do not follow from the commitments and entitlements associated with being either liberal or conservative in any logical (or nomological, for that matter!) way. Consequently, the unity of such categories is not explained by collective acceptance. Instead, we could say that the reasons for the clustering of just those properties are often unknown to the individuals participating in the reproduction of such kinds. I call this property of social categories *opaqueness*. Opaqueness does not mean that there are no relations between such properties. But it raises a challenge to the collective acceptance approach, because it suggests that its account of the multiple projectability of social categories is only partial.

### 4 Historical reproduction as an explanation of multiple projectability

Turning our attention to the processes reproducing social reality takes us back to Millikan’s account (see Sect. 1), as developed by Godman (2015, 2020). Millikan draws a distinction between eternal and historical kinds as substances with fundamentally different explanations for their multiple projectability. Some real kinds (like protons, or gold) are eternal in the sense that their property cluster is determined by laws of nature. They can be said to have an essence, an inner nature which by natural necessity results in a certain set of observable properties. Other kinds like dogs, peanut butter, and gothic cathedrals are historical in that there is a contingent cause of their clustering, and the similarities between their members are explained by spatiotemporal contiguity (Millikan, 2000, pp. 18–20).

More precisely, for historical kinds, the kind-characteristic cluster of properties is explained by the (1) shared environment between kind members and (2) processes of historical reproduction and transmission. Tigers resemble other tigers not because tigers occupy an atemporal slot in the order of reality (i.e. Porphyry’s tree) but because tigers form a reproductive lineage: Descendants inherit their properties from their parents through an (imperfect) process of copying and recombination. "It is not just that each exhibits the properties of the kind for the same ahistorical or eternal reason. Rather, each exhibits the properties of the kind because other members of the kind exhibit them" (Millikan, 2000, p. 20). This distinction between the two explanations for multiple projectability cuts deep, and it implies a deep modal difference between the two types of kinds: given the laws of nature, eternal kinds could not fail to carry the properties they in fact have, whereas different initial
conditions (for speciation, for instance) can lead to radically different properties of historical kinds.

Social categories obviously fall on the historical side of this divide, and Godman (2015, 2020) has applied the historical reproduction account to social categories such as social democracy, Christianity, and Sikhism. She treats these social kinds as cultural homologues and argues that that the historical continuity underlying them explains their inductive fecundity. Her account is not afflicted by the opaqueness problem. In it, Christians (or liberals in the US) share various properties not only due to their religious or political beliefs but also due to factors they may be oblivious about: teaching, social learning, imitation among kind members, and a shared environment. The empirical work by DellaPosta and colleagues dovetails nicely with this historical reproduction account: they propose interpersonal influence as a central explanation of lifestyle enclaves. By this, they mean processes very similar to ones Godman identifies: “Influence includes mechanisms that lead people to become more similar, through diffusion, contagion, imitation, assimilation, cooperation, convergent competition, or a host of other processes—or more dissimilar—through repulsion, divergent competition, differentiation” (DellaPosta et al., 2015, p. 1487).

The historical reproduction approach can account for the contingency of kind-characteristic property cluster: If in the 1960s, conservatives had had long hair, and in response, hippies had shaved theirs bald, it seems likely that these properties would have been passed on along cultural lineages. It may well be that many of the properties (re)produced along the historical continuum are unintended consequences of social action, and category members need not have (correct) understanding of why such properties are associated with their category.

However, I do not think shared history is either necessary or sufficient for multiply projectable social categories. First, necessity. The proponents of the historical reproduction approach state that being a part of a lineage is a necessary condition for kind membership (Godman et al., 2020; Millikan, 1998). This is why swamp tigers (see Godman, 2020, chapter 4) could not be real tigers, even if they shared all of their kind-specific properties with them. Notice that this is a metaphysical claim about kind membership, and a relatively tautological one at that: Membership in a historically determined kind requires belonging to the correct spatiotemporal lineage. But from the perspective of inductive inference, a kind consisting of both tigers and swamp tigers could be just as reliable as the one excluding swamp tigers, if the members of such a kind in fact shared their properties (excluding those logically related to the ancestry), and could even interbreed, as, I imagine, the thought experiment would suggest.

Next, sufficiency. It is certainly true that, by definition, processes of mutual copying and reproduction give rise to resemblances between the instances of a kind. But as proponents of the approach themselves acknowledge, copying is rarely perfect. Sometimes it leads to rich shared property profiles, but in other cases it does not. In

---

2 Just like the Swampman (Davidson, 1987), the swamp tiger was not born from its mother, but instead, it was created by a miraculous stroke of lightning in the marshes.
some cases, only a handful of properties might be passed on to new generations, or the mutation rate could be so high that the copying process fails to sustain a stable property cluster altogether. In biological contexts, different selection pressures may lead to either thick or thin, robust or fragile clusters of properties. Furthermore, cultural lineages range from highly conservative processes (e.g., book copying) to ones involving high levels of variation and improvisation (fashion).

Consequently, referring to processes of reproduction provides only a partial explanation of multiple projectability. More detailed attention on the nature of processes is needed to explain why some social categories seem to develop into recognizable lineages giving rise to substance-like kinds whereas others do not. History alone is not sufficient, as it brings about both stability and change.

5 Homeostatic mechanisms and entrenched social roles

An influential mechanism-based account of natural kinds was put forward by Richard Boyd (1991, 1999). In Boyd’s account, natural kinds have a bipartite structure. Their members share a cluster of properties, and a homeostatic mechanism governs the co-occurrence of properties in the cluster. I consider both the historical reproduction account and Boyd’s homeostatic property cluster account as reductive theories of kinds, because in both accounts, the existence of natural kinds is not a simple fact about reality. Instead, it can be explained by referring to the causal mechanisms governing the co-occurrence of the properties that characterize the kind.

For example, at standard temperature and pressure, the crystal structure of aluminum is a face-centered cube. This means that there is one atom on each corner of the cube and one at the center of each face. Roughly, the homeostatic property cluster account explains the observable properties of aluminum (its shininess, malleability, etc.) by referring to how the atoms are organized into a metal lattice. In this case, the metal lattice governs the co-occurrence of the properties of the kind, so we could view it as constituting the homeostatic mechanism of the kind. As another example, the biological species concept (Mayr, 1984) provides us with an account of what makes different biological species natural kinds: The homeostatic mechanism of the kind consists of the circulation of genetic material between the species members. Combined with reproductive isolation from other species, such circulation sustains the homeostasis of the species-specific genetic, morphological, and behavioral properties.

In a recent book, Ron Mallon (2016) applied Boyd’s mechanistic account to social categories (see Kuorikoski & Pöyhönen, 2012 for a similar account). On the face of it, Mallon’s social role constructionism seems quite compatible with Searle’s account: For a social role to exist, there must be (1) a representation, a label that picks out a category of persons and attaches them a description. Beliefs associated with the label must be (2) common knowledge (Lewis, 1969) in the community. However, as I suggested above, the existence of social roles in the community does not yet explain multiple projectability. It is only through the causal processes or mechanisms triggered by the existence of the social role that the category may
acquire multiple projectability. Mallon distinguishes between three kinds of mechanisms:

First, social roles may give rise to both non-strategic and strategic intentional actions. Non-strategic actions occur because people believe the description associated with the category. In contrast, strategic action is motivated by reasoning from the fact that the description is common knowledge, i.e., that others possess that knowledge as well. Say, for example, that I am a male in my early 20s. I come to know about research findings that suggest that men like me often tend to be aggressive (maybe due to hormone levels). Non-strategically, I might try to avoid threatening situations so as not to end up in circumstances where I could act irresponsibly. An example of strategic action would be joining a charity whose members help offer companionship to old people. Engaging in such activities would function as a credible signal of my peaceful intentions.

Secondly, social roles may generate automatic reactions. Work by Tajfel (1970) using the minimal group paradigm is a classic example of such mechanisms in action. Merely making people aware of labels and the existence of groups can have significant behavioral effects.

Third, Mallon (2016, p. 82) describes as “environmental construction” the processes where “common knowledge representations of human categories guide the modification of the environment in ways that systematically affect both members and nonmembers of the categories and also feed back into representational processes.” For example, our understanding of social categories gets reified in books, schools, museums, or public art. Common knowledge about kinds of people thereby influences our institutions, norms, and conventions that regulate behavior.

Non-strategic and strategic action, automatic reactions, and environmental construction form three different pathways through which members of social categories may obtain new properties. Significantly, in all three cases a kind emerges in representation-category interaction, where members of the kind react to common knowledge of the category, or to the practices or institutions informed by that knowledge. Such processes of representation—category interaction assume a temporal order where first a social role gets created, and then common knowledge concerning it induces behavioral effects. These changes in behavior may sometimes require us to revise our common knowledge regarding the category, and so on. This dynamic has been referred to as the looping effect of human kinds (Hacking, 1995a).

Without doubt, processes instantiating representation—category interaction constitute a significant class of mechanisms for generating multiple projectability. Returning to our example of political orientations in the US, Mallon’s theory appears to capture many of the processes of interpersonal influence discussed by DellaPosta and coauthors. For example, in-group imitation and differentiation from out-group members can be seen as instances of representation—category interaction: Despite my initial tastes in music or fine arts, I may try to acquire the habitus of a cultural omnivore characteristic of political liberals. Likewise, personal preferences for hunting or maintaining old cars may be suppressed due to anticipated disapproval by other in-groupers.

Such processes may not cover the whole range of relevant mechanisms. Based on their theoretical work, DellaPosta and coauthors suggest homophilous selection.
as another explanatory mechanism for the political ideology puzzle. Homophily refers to processes of selection, including self-selection, so that “ties of amity” are more likely to form between similar persons. Homophily presents the reverse temporal dynamic to interpersonal influence: shared properties between people may lead them to later share a group membership. In other words, people may seek to belong to groups with people who they resemble, and to avoid sharing labels with people who they regard as dissimilar. Consequently, the cluster of properties characteristic of a kind may (at least initially before processes of influence kick in) result from a process of sorting or filtering. I return to this possibility in the next section. Here I only wish to note that the homeostatic mechanisms described by Mallon’s account may not capture the whole range of relevant processes, and may therefore bias our conception of the grounds for the unity of social categories.

Before moving on, another remark. The discussion above suggests that the homeostatic property cluster theory does not help us distinguish between two roles that a mechanism may serve: generating inductive thickness and sustaining it. Such mechanisms are not always the same. The unstable and fleeting categories described by Ian Hacking (1995b, 1998) in his studies of looping effects provide evidence that some domains of social life feature influential, inductively rich, but temporally unstable phenomena (e.g., fashion, transient mental illness).

Consequently, the more we focus on the differences between the functions served by distinct mechanisms, the less work the homeostatic property cluster account itself seems to be doing. As with the historical reproduction account, when it comes to explaining the multiple projectability of social categories, the philosophical theory of kinds itself is relatively peripheral. In other words, our available philosophical accounts of social kinds (as real kinds) seem to provide answers to a rather limited set of questions concerning the inductive fecundity of social categories.

6 First boundaries, then entities

I now turn to a picture of social categories that has not, as far as I know, been articulated in the philosophical discussions on social kinds. In a 1995 article entitled “Things of boundaries”, sociologist Andrew Abbott describes the emergence of the role of social worker in the US at the turn of the twentieth century. Abbott shows how various activities conducted in hospitals, churches, and settlement houses in the 1870s would later be done by social workers. No particular group performed these activities, nor were these tasks understood as social work. But only a few decades later, there was training, professional societies, and journals for social workers. Previously separate tasks were brought together under the responsibility of the social worker, and altogether new tasks and concepts emerged, such as investigating family structures with an eye to their impact on health, as well as concern about stress in daily life.

This emergence of a new social category, a new kind of professional, the social worker, cannot be accounted for in terms of Searle’s collective acceptance account. Surely there were institutional speech acts establishing some members of the set of deontic powers characteristic of a social worker, and collective acceptance
concerning such assignments. But Searle’s account, and the historical reproduction account, cannot explain why those previously separate activities and tasks were brought together. Or why kindergartens and probation ultimately were not included in the social worker’s responsibilities, as was initially the case.

Nor can the set of properties characteristic of social workers (qua social workers) be fully, or even predominantly, explained by the representation—category dynamic explored by Mallon (2016). The various properties characteristic of social workers did not cluster only because people responded to being classified as social workers or to new environmental conditions in their work. Instead, Abbott portrays the emergence of the new institution as the result of what could be called boundary work (see Gieryn, 1983). For the most part, the elements constituting the profile of a social worker already existed in the social world, and the novel category emerged as a consequence of (de-)emphasizing certain differences between people, drawing boundaries in new ways, and combining such operations to “stitch” together various differences together into a category that had an inside and an outside. Through this process, various pre-existing properties were “recruited” into the category, thus making it both more real and inductively rich.

Abbott (1995, p. 860) writes: "Social entities come into existence when social actors tie social boundaries together in certain ways. Boundaries come first, then entities." Although metaphorical, I believe this emphasis on boundary formation and maintenance brings to light an important aspect of the constitution of social reality. Instead of starting from pre-existing categories or kinds, interactionist social theory (see Emirbayer, 1997) examines how social reality emerges from processes of interaction.

Focusing on boundaries provides an alternative explanation of how multiple projectability can come about. The approaches discussed above in Sects. 3–5 portray multiple projectability as involving the generation of novel properties: new deontic powers by fiat, copying from previous generations, and reactions to a classificatory label giving rise to new often unanticipated behaviors. Abbott’s account, instead, begins from the picture of a population of differences between individuals. We resemble, and differ from, each other in myriad ways. This landscape of differences can be cut into categories in various ways, depending on how boundaries are drawn.3 Abbott describes this process of (re)organizing a pre-existing “field” of things that as follows: "the making of an entity is simply the connecting up of these local oppositions and differences into a single whole that has a quality which I shall call ‘thingness’". It is this often overlooked aspect of social construction that I refer to as “recruitment.”

In the interactionist tradition, Tilly (2004) and McAdam provide more concrete examples of the mechanisms of boundary formation and boundary change. In their study of contentious political action, McAdam et al. (2001, pp. 140–148) distinguish between three mechanisms of category formation:

---

3 Bach points out that Haslanger (2012, p.188) relies on a similar metaphysics of plurality in her attempt to motivate the role of background values in the classification of humans.
Invention involves authoritative drawing of new boundaries and prescription of relations across the boundary. The example given by McAdam and coauthors is Bosnian Serb leaders deciding who counts as a Serb, and how Serbs are supposed to interact with non-Serbs.

Borrowing is process where an existing combination of boundaries and relations gets installed in a new setting. One could imagine that the liberal—conservative division introduced in Sect. 2 could be imported to new contexts to structure a wide range of social settings that were initially independent from political dividing lines.

Encounter refers to a process where members of previously separate but internally connected networks come into contact with each other, begin competing for resources, and interactively negotiate definitions of a boundary and the relations across it. Again, the process involves not so much the generation of new thick clusters of category-specific properties as emphasizing and legitimizing pre-existing differences between categories. As an example, the 2015 European refugee crisis surely involved the accentuation of category boundaries through encounters between people seeking refuge and living in Europe.

The mechanisms of invention, borrowing, and encounter involve aspects that can be captured by the collective acceptance theory or Mallon’s mechanisms of representation–category interaction. What sets the McAdam’s mechanisms apart, however, is the assumption of a pre-existing field of differences, where processes of strategic boundary formation introduce order and change. Likewise, homophilous sorting between liberals and conservatives exemplifies non-strategic boundary work conducted by the category members themselves. Although none of these mechanisms guarantees the emergence of multiply projectable social kinds or categories, they exemplify a possibility often ignored in the philosophical literature on social categories. In these cases, inductively rich categories are not generated ex nihilo but instead they throw property profiles into relief through boundary work that legitimizes or delegitimizes some differences at the cost of others, thereby giving rise to novel inductive affordances.

7 Social categories, real kinds?

I conclude with four observations regarding the status of social categories as real kinds. First, finiteness. Mill (1891/2001) distinguished real kinds from finite classes in that the members of real kinds share innumerably many properties, whereas the commonalities between members of finite classes are exhausted by linguistic conventions. As the discussion on mechanisms of boundary formation suggests, such radical inductive fecundity will not apply to most social categories. Instead, kinds assembled by redrawing category boundaries and thus recruiting properties into the category need not have an extensive set of projectable properties. Therefore, although social categories may be multiply projectable, the finite size of the set of shared properties distinguishes them from eternal kinds, as defined by Millikan. However, as Hacking (1991) has observed, it is not clear that even our most typical examples of kinds in the natural sciences satisfy Mill’s inexhaustibility criterion.
Another obvious difference between social and eternal kinds has to do with explanations of stability. Let us briefly return to the distinction between mechanisms of category formation and maintenance. In Sect. 1, two dimensions of multiple projectability were distinguished: the range of inferences and their reliability. Bringing the issue of diachronic category stability into view suggests a third dimension, stability of over time: For all practical purposes, multiple projectability of a category seems to require that the inferences concerning it are (to some degree) reliable over time. The mechanisms of category formation discussed above do not for the most part explain such stability: hence, it requires a separate explanation. Moreover, cases of the looping effect provide empirical evidence that such stability should not be taken for granted. I do not believe that this distinction applies to eternal kinds: For example, no separate explanation could meaningfully be required for the emergence of the properties of gold and the stability of the property set. If laws of nature guarantee gold having the properties it in fact has, the invariance of laws over time also explains the persistence of the property set.

The third observation concerns the compatibility of nominalism and multiple projectability. Abbott (1995, p. 860) believes our thinking about the social world is distorted by a substantivist bias: “We think about social entities as overgrown versions of […] biological individuals, and, thus, have become accustomed to think that social entities have essences like biological individuals, that they have some internal plan or thingness or Aristotelian substance.” I think this poverty-of-organicism argument does diagnose a problematic aspect of much of the philosophical thinking around social categories. Our philosophical accounts of kinds tend to reflect a “reductionist heuristic” (Wimsatt, 2007) where we often isolate a category from its context and analyze its properties in isolation. Such an approach easily leads to substantivist thinking. The relational approach suggests, instead, that apparent “thingness” of (purely) social categories could be compared to a standing wave in a medium, the coincidence of several boundaries giving rise to an impression of a substance or a solid entity. This seemingly strong nominalism appears, however, quite justified, especially in the case of social categories. Although categories such as ‘liberal’ or ‘conservative’ can serve as substance concepts in the sense that (1) they organize our conceptual ecology and (2) (re)identifying their members triggers and justifies inferences about various other properties, it would seem absurd to posit that the ontological asymmetry between a kind and a property would be manifested between, say, ‘conservative’ and ‘bird-hunter’. Though a conservative person in the forest may “carry” the property of being a bird-hunter, there surely is no ontological asymmetry between the referents of ‘conservative’ and ‘bird-hunter’ as such; it is not the case that one is a kind and the other a “mere” property. Rather, for some fully contingent reason, social mechanisms of boundary formation, reproduction, and representation–category interaction have functioned around a focal property, which in this case happened to be political orientation. Such functioning led to a network of correlations where political orientation emerges as a hub. Consequently, two kinds of realism about social categories must be kept separate. To treat some social categories as substance concepts and to assert that they support multiple projectability, just like natural kinds, one need not make commitments to the ontological statement that kinds (over and above their members) exist as substance-like entities.
Finally, a point about reality, acceptability, and intervention. Claims about the social constructedness of a phenomenon should not be seen as implying eliminativism concerning that phenomenon. Instead, they aim to show its contingency, that what we thought was natural and immutable may by artificial and changeable (Hacking, 1999). Nonetheless, such categories may be real in the sense of being multiply projectable and playing an important role in the mental and social lives of the people involved. We may find such categories unacceptable, for example, if they embody prejudice and discrimination. Having a clear view of the forces that explain the emergence and stability of the category allows us to aim our moral evaluation at the right kinds of targets. For example, uncovering the strategic boundary work underlying unacceptable racial or ethnic categories allows us both to explain, and to attribute blame to the processes that construct and sustain them. Likewise for interventions: In order to be effective, ameliorative conceptual interventions (e.g., Haslanger, 2020) should be based on knowledge of the causal processes involved (Bach, 2019). As argued in this article, not all such processes manifest the representation—category dynamic characteristic of looping effects, and, consequently, may not be altered by debunking strategies such as simply revealing their existence (cf. Mannheim, 1952). Furthermore, an ameliorative intervention itself could be seen as an instance of boundary work affecting symbolic boundaries, conducted not by members of the in- or out-group themselves, but instead originating from the epistemically authoritative position of a researcher.

8 Conclusion

In this article, I have formulated a puzzle concerning the multiple projectability of social categories. As a point of departure, I examined Searle’s collective acceptance account of social institutions. Although it does take the first step to understanding how social categories can begin to acquire inductive richness, it fails to account for the opaque and thick property clusters characteristic of many social categories. Millikan’s and Godman’s historical reproduction approach usefully emphasizes the role of learning, imitation, and copying as mechanisms that reproduce category-specific properties over time. It cannot be the whole story, however, because it fails to explain differences between historical processes that bring about resemblance and those that give rise to variation. Mallon’s social role constructivism provides an account of the feedback mechanisms between the representation of a category and the behavior of its members, which can lead to the emergence of novel properties for those members. Yet, the processes involving boundary formation and “recruitment” reverse the temporal logic of Mallon’s mechanisms. More generally, I suggested that approaching social categories from the perspective of boundary formation exposes a potentially misleading root metaphor in how we often think about social kinds and social construction. I propose an alternative starting point that leads to conceptualizing social kinds not as self-subsisting entities, but as often fleeting, though sometimes stabilized, outcomes of boundary-drawing.

Together with the accounts already presented in the literature, the recruitment account takes us one step closer to explaining multiple projectability. At the same
time, boundary-drawing reflects a population-thinking approach to the social world: we differ from each other in numerous ways, only some of which are deemed as relevant. Which ones we end up emphasizing may lead to emergence of different, even cross-cutting categorizations of social phenomena and kinds of people.

Acknowledgements I am grateful to the two anonymous reviewers, as well as the audiences at Constructing human kinds workshop (University of Copenhagen, 2018), Social Ontology 2019 conference (Tampere University), Perspectives on Science seminar (University of Helsinki, 2021), and Reactivity and Categorization workshop (University of Copenhagen, 2021) for constructive feedback on earlier versions of the paper, and to Kate Sotejeff-Wilson for language editing.

Funding Open access funding provided by University of Helsinki including Helsinki University Central Hospital. The research was funded by the Academy of Finland, Application Number # 336324.

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

Abbott, A. (1995). Things of boundaries. Social Research, 857–882.
Bach, T. (2019). Real kinds in real time: On responsible social modeling. The Monist, 102(2), 236–258.
Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. Philosophical Studies, 61, 127–148.
Boyd, R. (1999). Kinds as the ‘workmanship of men’. In J. Nida-Rümelin (Ed.), Rationalität, realismus, revision (pp. 52–89). Walter de Gruyter.
Craver, C. (2004). Dissociable realization and kind splitting. Philosophy of Science, 71, 960–971.
Davidson, D. (1987). Knowing one’s own mind. Proceedings and Addresses of the American Philosophical Association, 60(3), 441–458.
DellaPosta, D., Shi, Y., & Macy, M. (2015). Why do liberals drink lattes? American Journal of Sociology, 120(5), 1473–1511.
Emirbayer, M. (1997). Manifesto for a relational sociology. American Journal of Sociology, 103(2), 281–317.
Gieryn, T. (1983). Boundary-work and the demarcation of science from son-science: Strains and interests in professional ideologies of scientists. American Sociological Review, 48, 781–795.
Godman, M. (2015). The special science dilemma and how culture solves it. Australasian Journal of Philosophy, 93(3), 491–508.
Godman, M. (2020). The Epistemology and Morality of Human Kinds. Routledge.
Godman, M., Mallozzi, A., & Papineau, D. (2020). Essential properties are super-explanatory: Taming metaphysical modality. Journal of the American Philosophical Association, 3, 1–19.
Hacking, I. (1991). A tradition of natural kinds. Philosophical Studies, 61, 109–126.
Hacking, I. (1995a). The looping effects of human kinds. In D. Sperber, D. Premack, & J. Premack (Eds.), Causal cognition: A multidisciplinary debate (pp. 351–94). Clarendon Press.
Hacking, I. (1995b). Rewriting the soul: Multiple personality and the sciences of memory. Princeton: Princeton University Press.
Hacking, I. (1998). Mad travelers: Reflections on the reality of transient mental illnesses. University of Virginia Press.
Hacking, I. (1999). The social construction of what? Harvard University Press.
Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. Vintage.
Haslanger, S. (2012). *Resisting reality: Social construction and social critique*. Oxford University Press.
Haslanger, S. (2020). Going on, not in the same way. *Conceptual engineering and conceptual ethics*. https://doi.org/10.1093/os/9780198801856.003.0012
Hempel, C. G. (1964). *Fundamentals of concept formation in empirical science*. University of Chicago Press.
Kuorikoski, J., & Pöyhönen, S. (2012). Looping kinds and social mechanisms. *Sociological Theory, 30*(3), 187–205.
Lewis, D. (1969). *Convention: A philosophical study*. Basil Blackwell.
Machery, E. (2009). *Doing without concepts*. Oxford University Press.
Mallon, R. (2016). *The construction of human kinds*. Oxford University Press.
Mannheim, K. (1952). Das problem einer soziologie des wissens. In K. Mannheim (Ed.), *Essays on the sociology of knowledge*. Routledge Kegan Paul.
Mayr, E. (1984). Species concepts and their application. In E. Sober (Ed.), *Conceptual issues in evolutionary biology: An anthology* (pp. 531–40). MIT Press.
McAdam, D. T., & Tarrow, S. (2002). S., & Tilly, C. (2001). *Dynamics of contention*. Cambridge University Press.
Mill, J. S. (1891/2002). *A system of logic: Ratiocinative and inductive*. University Press of the Pacific.
Millikan, R. G. (2000). *On clear and confused ideas: An essay about substance concepts*. Cambridge University Press.
Millikan, R. G. (1998). A common structure for concepts of individuals, stuffs, and real kinds: More mama, more milk, and more mouse. *Behavioral and Brain Sciences, 21*(1), 55–65.
Murphy, D. (2006). *Psychiatry in the scientific image*. MIT Press.
Pöyhönen, S. (2014). Natural kinds and concept eliminativism. In V. Karakostas & D. Dieks (Eds.), *EPSA11 Perspectives and foundational problems in philosophy of science* (Vol. 2, pp. 167–179). The European Philosophy of Science Association Proceedings.
Searle, J. (1995). *The construction of social reality*. Free Press: London.
Sveinsdóttir, Á. K. (2013). The social construction of human kinds. *Hypatia, 28*(4), 716–732.
Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American, 223*, 96–102.
Tilly, C. (2004). Social boundary mechanisms. *Philosophy of the Social Sciences, 34*(2), 211–236.
Wimsatt, W. (2007). *Re-engineering philosophy for limited beings: Piecewise approximations to reality*. Harvard University Press.

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