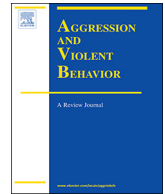




Contents lists available at ScienceDirect

Aggression and Violent Behavior

journal homepage: www.elsevier.com/locate/aggiobeh

Classification, kinds, taxonomic stability and conceptual change

Jaipreet Mattu^{a,b}, Jacqueline A. Sullivan^{a,b,c,*},¹^a Department of Philosophy, University of Western Ontario, 7170 Western Interdisciplinary Research Building, 1151 Richmond St., London N6A 5B8, Ontario, Canada^b Rotman Institute of Philosophy, University of Western Ontario, 7170 Western Interdisciplinary Research Building, 1151 Richmond St., London N6A 5B8, Ontario, Canada^c Brain and Mind Institute, University of Western Ontario, 7170 Western Interdisciplinary Research Building, 1151 Richmond St., London N6A 5B8, Ontario, Canada

ARTICLE INFO

Keywords:

Concepts
Construct stability
Classification
Natural kinds
Philosophy of science

ABSTRACT

Scientists represent their world, grouping and organizing phenomena into classes by means of concepts. Philosophers of science have historically been interested in the nature of these concepts, the criteria that inform their application and the nature of the kinds that the concepts individuate. They also have sought to understand whether and how different systems of classification are related and more recently, how investigative practices shape conceptual development and change. Our aim in this paper is to provide a critical overview of some of the key developments in this philosophical literature and identify some interesting issues it raises about the prospects of the so-called “special sciences”, including psychiatry, psychology, and the mind-brain sciences more generally, to discover natural kinds.

1. Introduction

Scientists represent their world, grouping and organizing phenomena into classes by means of concepts. Philosophers of science have historically been interested in the nature of these concepts, the criteria that inform their application and the nature of the kinds that the concepts designate. They also have sought to understand whether and how different systems of classification are related and more recently, how investigative practices shape conceptual development and change (e.g., Feest & Steinle, 2012; Kendig, 2016a). Our aims in this paper are to provide a critical overview of some of the key developments in this philosophical literature and identify some interesting issues it raises about the prospects of the so-called “special sciences”, including psychiatry, psychology, and the mind-brain sciences more generally, to discover natural kinds.

We begin, in Section 2, with an historical overview of philosophical thinking about classification. Although philosophers and scientists generally agree that the aims of classification are broadly epistemic, they disagree about the nature of the kinds of things the world contains, the appropriate methods for individuating kinds and grouping them into categories and the relationship between the resulting classification systems and the world. One overarching aim of scientific classification is the discovery of what philosophers refer to as “natural kinds”, and we move on in Section 3 to consider different understandings of this

concept. We explain that debates about whether different areas of science are able to achieve “the natural kinds ideal” hinge not only on how we define the concept of a natural kind, but also on the stability of the phenomena under study in the areas of science being considered. It also depends crucially, as we explain in Section 4, on the conceptual and methodological practices of investigators in those areas of science.

2. Principles of classification: a brief historical overview

2.1. Conventionalism

Human beings impose conceptual order on their world. In the most basic terms, we are born into a world of language users; learning the meanings of words requires the abilities to detect objects having certain properties in the world, to recognize similarities and differences among those objects with respect to those properties and to understand that things that others identify using the same name share certain properties in common. We learn these basic rules of classification and assign things that share properties in common to groups having unique names. Common examples of groups include stars, mammals, trees, beliefs and feelings. We learn that stars have the basic properties of being luminous and visible in the night sky and that beliefs are things that humans and some non-human animals have that can be assigned a truth value. Sometimes we place unlike things into the same group—we *misclassify*

* Corresponding author at: Department of Philosophy, University of Western Ontario, 7170 Western Interdisciplinary Research Building, 1151 Richmond St., London N6A 5B8, Ontario, Canada.

E-mail address: jsulli29@uwo.ca (J.A. Sullivan).

¹ Co-authors had equivalent input and are listed in alphabetical order.

<https://doi.org/10.1016/j.avb.2020.101477>

Received 30 May 2020; Received in revised form 6 July 2020; Accepted 6 July 2020

1359-1789/ © 2020 Elsevier Ltd. All rights reserved.

either accidentally or intentionally (perhaps, in some cases, to get clear about the rules).

The concepts that we learn—that ordinary people use to classify the world—are sometimes referred to as “folk” concepts (e.g., Churchland, 1981). These categories are so-called because while they enable ordinary language users to effectively navigate the day-to-day world, they arise from convention. We learn how to use concepts and we may use them without explicitly questioning what they mean or what the things that we group similarly actually share in common beyond what we can detect by means of our senses. Folk concepts and classification systems based exclusively on convention are widely regarded as insufficient stopping points for classifying our world if our aims are explanation, prediction, intervention and control. Yet, what principles should guide the development of a classification system? In this section, we consider some of the key answers to this question on offer in the philosophical literature.

2.2. Essentialism

One of the earliest answers to how to properly classify things in the world that remains influential today originated with the Ancient Greek philosophers. In the *Categories*, Aristotle develops a system of classification rooted in a doctrine known as *essentialism*, which has its origins in Platonic thought. According to Plato, we exist in a world of sensible objects, but our senses do not reveal the true nature or essence—the *Forms*—of those objects. In order to move beyond the realm of sensory appearances and grasp the essential nature of things, we must engage in the method of dialectic (Grube, 1974, 165), and evaluate particular cases of things in the world in order to abstract the fundamental properties they share in common. Aristotle explicitly engages in this method in the *Categories* (Ackrill, 1963), putting forward a set of category names (i.e., concepts) and providing definitions for them in terms of properties that inhere in and are essential to the things to which those names apply (i.e., essences). For example, he specifies what it is to be a *man* – what the essential properties of being human are—by distinguishing those properties (e.g., being an animal) from ones that are typically ascribed to individual men (e.g. proper names) or accidentally associated with the concept “man” (e.g., two-footed).² Each Aristotelian category constitutes a set of severally necessary and jointly sufficient conditions for a thing to be a member of that category (e.g., Hull, 1965).

Several features of Ancient essentialism are worthy of note. First, the method of specifying necessary and sufficient conditions for category membership aims to advance classification beyond ordinary folk understandings of the nature of things. Plato was explicit that grasping the underlying reality of nature was something that only dialectically trained philosophers could do. A second noteworthy feature is that once a set of necessary and sufficient conditions for category membership is identified, the resulting conceptual category is intended to be stable; by engaging in dialectic, one arrives at the *correct* classification system for understanding the true nature of reality. Plato’s *Forms* and Aristotle’s *Categories* were the endpoints of dialectical inquiry, and they were understood to be unchanging foundations of knowledge akin to the axioms and definitions of geometry and mathematics (e.g., Williams, 2001).³

²The examples of categories that Aristotle puts forward in his treatise are of more fundamental types, such as *substance* and *quantity*. However, throughout the text he uses the example of “man” to illustrate what features are essential to being human and which features may be accidentally ascribed that are not essential to how the concept of man is defined.

³According to philosopher Michael Williams (2001), the Ancient Greeks believed that knowledge once arrived at was stable. The model for such stability was mathematics; from a set of first principles that included axioms and definitions one could deduce other true claims by means of mathematical proof. The Ancient Greeks believed that just so long as they could identify similarly

2.3. Empiricism

The British empiricist, John Locke, put forward a different basis for an essentialist classification system. Locke noted that the objects of our everyday experience have properties detectable by the senses and we use those properties to differentiate objects from each other and place them into groups to which we apply specific names. He referred to groupings based on sensible properties of objects and shaped by social convention as “nominal essences” and differentiated them from hidden microstructural properties that he believed constituted the “real” essences of things (Locke, 1689[1975], 417). Locke explains, for example, that although we customarily associate a name like “gold” with instances of metal having a specific color, texture and density, these features cannot be used to individuate the real underlying microstructural properties on which these sensible properties depend (Locke, 1689[1975], 419).

As an empiricist, Locke believed that the scientific method was fundamental for advancing human understanding beyond “nominal essences” towards classification systems based on “real essences”. Yet, he was skeptical that science could illuminate the microstructural properties of objects and how those properties mechanistically give rise to “nominal essences”. Insofar as he described the material world as being in a constant state of flux, he seemed open to the possibility that if science were to advance, classification systems informed by a knowledge of the “real essences” of things may be subject to on-going revision (cf. Locke, 1689[1975], 419).

The Scientific Revolution prompted later empiricists, including William Whewell and John Stuart Mill, to be more optimistic about the prospects of science for discovering “real essences” and developing “real” classification systems. By the 18th century, classification systems had been developed in sciences like botany, zoology and mineralogy. Criticisms of the “naturalness” of those classification systems and proposed revisions from scientists working in these scientific domains prompted Whewell and Mill to try to clarify the principles that should inform scientific classification. At that time, philosophers of science regarded Newtonian mechanics as exemplifying the proper method for how to do science (see Herschel, 1830 [2009]; Whewell, 1840 [2014]). It was widely accepted that Sir Isaac Newton had employed the method of induction; he started with the observation of many cases of phenomena to determine the important features that like cases shared in common; he then used those observations as a basis for inferring causes or mechanisms, arriving at a set of laws that could explain a wide range of mechanical phenomena.

Whewell (1840 [2014]) acknowledged that different systems of classification may serve different purposes. For example, in his day, whales were classified by whalers as fish, but given that female whales (cows) had mammary glands and nursed their young, they clearly were not “fish”, but rather, “suckling beasts or mammals” (1840 [2014], 456). Scientists, according to Whewell, should aim to go beyond classification systems that are characteristic of “common practical life” and

(footnote continued)

secure foundations for philosophy, logical proofs could be used to deduce other true claims, and those claims would be as secure as were their foundations. With the rise in the use of the inductive method following the Scientific Revolution, however, and the fact that inductive inferences are ampliative because the claims arrived at by means of induction go beyond what the premises actually support, there was a growing awareness that the foundations of scientific knowledge are not stable. Moreover, there was a growing appreciation of the fallibility of science that emerged prior to (e.g., Bacon, 2000), during and after the Scientific Revolution consistent with the idea that the scientific method can sometimes fail, and scientific claims are indeed revisable in light of new information. Indeed, as we explain later in this section, John Stuart Mill, insofar as he claimed that the discovery of “real kinds” will be unending in science, seems to have been amenable to the idea that scientific concepts and classification systems will never be stable.

identify “coherent and systematic collection[s] of properties” that constitute “Natural Systems” of classification that individuate the “Natural Affinit[ies]” of things (1840 [2014], 470).

Whewell, however, noted that not every classification system produced by science will necessarily be “natural”. Scientists themselves may fail to recognize fundamental properties that ideally ought to inform the development of a classification system. This may happen due to improper use of the inductive method or technological limitations. Whewell provides the example of Swedish botanist Carl Linnaeus’s (1707–1778) taxonomy of plants, which divided them into groups based on pistil and stamen number. Whewell remarks that such artificial characteristics are used as a basis for classification primarily when natural relationships that hold between things are not yet known, and artificial classifications are subject to revision as natural relationships are illuminated. Noting the reasons botanists provided for rejecting Linnaeus’s classification system, Whewell remarks that “it is plain that they seek something, not of their own devising and creating; —not anything merely conventional and systematic; but something which they conceive to exist in the relations of the plants themselves; —something which is without the mind, not within; —in nature, not in art; —in short, a natural order” (1840 [2014], 474). Artificial classification systems like that of Linnaeus, could, however, according to Whewell, provide roadmaps for natural classification systems by “plac[ing] us in a situation where the detail is within our reach” (1840 [2014]480).⁴

Mill (1874), like Whewell, regarded the process of developing a classification system as tentative and proceeding on a case-by-case basis in which two things are compared and then more things are added to the comparative process as a means to determine their similarities and differences. This tentative process is supposed to be inductive and facilitate the discovery of “general truths, that is, truths applicable to classes” (Mill, 1874, 807). According to Mill, only important properties of things that may correspond to laws ought to be grouped together. For example, we may look out at the animal kingdom and decide to classify animals on the basis of their color; so, we place black bears and jaguars in one group and seals and brown bears in another. Yet, such groupings are not, from Mill’s perspective, important, or “real” because they do not tell us much about other properties of animals—color is not connected to more fundamental features of animals, such as their physical structures or their behaviors. Mill thought that in order for a classification system to pick out “real” as opposed to “artificial” kinds, the groupings ought to home in on similarities and differences connected to “many other important particulars” of the kinds of things being classified (1847, 802). Additionally, Mill believed that the process of discovering “real kinds” – i.e., identifying the important properties of kinds and developing classification systems that individuate “real kinds”, is ongoing in science, which is consistent with the idea that classification systems and scientific knowledge more generally are not stable.⁵

2.4. Logical empiricism

Logical empiricists in the 20th century also were interested in the foundations of scientific classification and the relationships between different taxonomic frameworks, theoretical terms and kinds under study in different areas of science. For example, in his paper *Fundamentals of Taxonomy* (1959/1965), Carl Hempel sought to specify the logical and methodological foundations of classification in empirical science and to tease out a set of implications for psychiatric classification. Hempel claimed that scientific classification systems go through at least two discrete stages of development. His understanding of the first, what he referred to as a “descriptive stage” of scientific

classification, was informed in part by the doctrine of operationalism advocated by the physicist Percy Williams Bridgman (e.g., [Bridgman 1927, 1938, 1952](#)). In this stage, according to Hempel, scientists aim to specify uniform and publicly observable testing operations for scientific terms, so that a given term may only be applied if a given test yields a specified outcome. For example, if psychologists want consistent and publicly verifiable criteria of application for a psychological term like “stress”, they might develop a test like the *Trier Social Stress Test*, and then specify the kinds of observable behaviors, including sweaty palms and increased heart rate, which must be elicited by a subject during the test in order to warrant application of the term “stressed” to that subject. Hempel claimed that this first stage of classification is “descriptive”, insofar as scientists must appeal to observable “surface” features of phenomena in order to operationally define concepts and place them into groups (see also [Chang, 2019](#); [Feest, 2005](#)).

Operationally defined categories were intended to facilitate communication among scientists having different theoretical perspectives who were “engaged in a common research project” but lack a shared vocabulary. Yet they served only as an important practical starting point for the development of scientific classification systems. As those empiricists before him, Hempel claimed that sciences should strive to move away from taxonomies based on “observable” features of phenomena to conceptual taxonomies having “systematic or theoretical import” ([Hempel, 1965, 146](#)). While he agreed with Mill that scientific categories must identify “important” and extensive clusters of properties that had a “high probability” of being associated with each other, he also thought they must reflect underlying regularities in the subject matter in a given scientific domain that could be expressed in terms of laws and general theories that facilitated explanation, prediction and understanding ([Hempel, 1965, 146](#)). Discovering such general laws or theories would prompt revisions to scientific taxonomies such that the categories would ultimately correspond to bona fide divisions in the natural world, or “natural kinds”.

Hempel puts forward classification in biology as an example of the kind of historical shift from descriptive to theoretical classification he has in mind. Taxonomies in biology originally classified organisms with respect to observable and predominantly morphological features. However, morphology is no guide to natural divisions in the biological order, and in response to the development of the theory of evolution by natural selection, these early taxonomies were replaced by taxonomies based on phylogenetic relationships ([Hempel, 1965, 147](#)). If we consider instead classification in psychology and psychological concepts like “stress”, “memory”, “attention” and “fear”, that have been historically defined in terms of observable changes in behavior in response to specific stimuli or tests, it is possible that advances in a different area of science, namely, neuroscience, may illuminate neurophysiological or other underlying factors that may prompt their replacement. However, it remains an open question what happens to concepts in science and entire systems of classification as more is learned about underlying causal structures or mechanisms. As we explain in the next section, there are good grounds for thinking that there is no one size fits all model for scientific classification and conceptual change in science, nor are the kinds under study in different areas of science subject to the same kinds of conceptual challenges.⁶

3. The natural kinds ideal

Even a brief analysis of historical views of classification like that provided in [Section 2](#) reveals a general consensus that if the aims of classification are broadly epistemic—to grasp the underlying structure of reality (i.e., beneath appearances) and/or, to explain, predict and

⁴ For further discuss of Whewell on scientific classification, see [Quinn \(2017\)](#).

⁵ For further discussion of Mill, see [Khalidi \(2013\)](#); [Magnus \(2012, 2014\)](#).

⁶ See for example, [Andersen \(2010\)](#); [Bloch-Mullins \(2020a, 2020b\)](#); [Brigandt \(2003\)](#); [Ereshefsky \(2004\)](#); [Feest and Steinle \(2012\)](#); [Franklin-Hall \(2015\)](#); [Griffiths \(2004\)](#); [Khalidi \(2013\)](#); [LaPorte \(2004\)](#); [Nersessian \(2008\)](#).

control, then the kinds a classification system individuates should be “real” or “natural”. While philosophers generally agree that science aims to discover *natural kinds*, they disagree about how to conceive of such kinds and respond differently to the question of whether science is capable of discovering them (see Bird & Tobin, 2018). In 20th century philosophy, a lot of ink has been spilt aiming to clarify what kinds of kinds there are and debating the metaphysical, ontological, and epistemological statuses of these kinds. In this section, we focus on a representative subset of natural kind concepts and consider arguments for and against the idea that sciences like biology and “special” or “human” sciences like psychology and psychiatry, can develop classification systems that track natural kinds.

Natural kinds are understood to share a basic set of properties in common. First, they are things that exist naturally in the world and mind-independently of us; their existence does not hinge on our ability to conceptually individuate them and they existed prior to our discovering them. Plato's Forms, Locke's microstructural properties, Whewell's natural affinities, Mill's “real kind” are all regarded by their proponents as real and (with Plato's Forms as the exception), bona fide parts of the natural order. Second, natural kinds share a core set of properties in common; these properties may be at the microstructural level, but they do not have to be (see e.g., Kincaid and Sullivan, 2014; Borsboom et al., 2018; Boyd, 1991, 1999; Kendler, Zachar, & Craver, 2011; Khalidi, 2013; Magnus, 2012; Mellor, 1977; Tabb & Schaffner, 2017). It also is widely assumed that once a conceptual category picks out a “natural kind” it permits scientists to make inductive inferences, formulate generalizations about that kind and discover other natural kinds (see Boyd, 1999; Kendig, 2016a, 2016b; Khalidi, 2013; Magnus, 2012). To take a paradigmatic example of a natural kind from chemistry, members of the category ‘chemical element’ are separated from non-members in terms of a single defining feature—their atomic structure or number, i.e., the number of protons found in the nucleus of every atom of a given element. Each element has a unique atomic number and thus, to say that chemical elements are natural kinds is to say that there is a *natural difference* that sets them apart from other things. Depending on their intuitions, philosophers of science make claims about the “naturalness” of the kinds that are discoverable in different areas of science from the perspectives of realism, constructivism, conventionalism, pragmatism, and pluralism. In this section, we will consider each of these perspectives in turn.

3.1. Essentialist kinds

Essentialists are realists about natural kinds. Plato and Aristotle believed that things in the world have *essences* that are detectable when one engages in directed consideration of similarities and differences among things in the world and identifies the fundamental properties—i.e., the essence of those things (e.g., Ackrill, 1963; Bird, 2010; Devitt, 2008; Ellis, 2001, 2008; Hull, 1965). On this Classical essentialist view, then, natural kind categories are groupings that consist of members that are said to share an *essence* that exists behind the world of appearances. Locke's essentialism, in contrast, is consistent with the idea that an essence is a set of intrinsic, fixed properties that underlie and cause observable features unique to a given kind. Modern philosophers often construe essentialism as a mixture of Classical and Empirical versions in claiming that essences have “necessary and sufficient properties that identify the member of that kind and from which all identifying characteristics of that kind arise in all members of that category” (Kendler et al., 2011, 1144). Proponents of the essentialist view put forward chemical elements as a standard example to defend their view. For instance, gold can be correctly identified solely on the basis of its atomic number of 79. Its atomic structure gives rise to unique observable features, such as its color, malleability, and melting point, making it easily distinguishable (Devitt, 2008; Ellis, 2001).

However, some philosophers of biology argue that the essentialist's assumption that a single essence can give rise to the properties

attributable to members of a kind is ill-suited for biological phenomena (Boyd, 1991, 1999; Hull, 1965). David Hull (1965), for example, criticizes the idea that a stable set of jointly sufficient and severally necessary conditions for category membership could be provided for each species given that their features change throughout evolutionary history. Boyd (1991, 1999) argues that biological species cannot be said to share an essence due to evolutionary processes such as mate selection and genetic drift. Genetic and environmental factors also give rise to different phenotypes among individuals within a given species population (Boyd, 1999).

Boyd's arguments against essentialism with respect to biological kinds have been extended to models of disease. While genetic disorders result from a single causal factor, other kinds of diseases involve a number of causal factors, spanning from the molecular to the physiological, anatomical, behavioral and environmental levels. For example, psychiatric research reveals that a given mental disorder can manifest through a spectrum of symptoms from person to person, suggesting that such disorders do not arise from a single cause (Kendler et al., 2011). As Zachar (2015, 289) notes, a person's symptoms may “evolve over time...certain symptoms coming into foreground, and then receding into the background as other symptoms take their place” and “interactions between symptoms can also generate new symptoms.” Answering to the diversity and instability of biological kinds are perhaps the most challenging criticisms that essentialists face, because they are in conflict with the aim of developing a system of classification based on a stable set of necessary and sufficient conditions for categorizing natural kinds.

3.2. Homeostatic property clusters

As an alternative to essentialism, Richard Boyd suggests a realist but non-essentialist view of natural kinds. He introduces the ‘homeostatic property cluster theory’ (HPC-Theory), through which natural kinds exist “not by any set of necessary and sufficient conditions, but instead by a ‘homeostatically’ sustained clustering of properties or relations” (Boyd, 1999, 141). On this account, natural kind categories consist of members that share co-occurring properties that reliably cluster together due to shared underlying causal mechanisms that sustain the property clusters. The HPC view is supposed to sidestep the issue that critics have associated with essentialism failing to accommodate the diversity and adaptability of biological kinds. Boyd (1999) remarks that it is not necessary for members of an HPC kind category to share every single trait, leaving room for variation between members of that kind. Boyd (1991, 142) considers biological species to be a paradigmatic example of HPC kinds insofar as a set of “imperfectly shared and homeostatically related morphological, physiological and behavioural features [...] characterize [...] members” of a species. To take an example, “certain anatomical structures, body type, and predatory behaviours form a homeostatic property cluster called ‘tiger’” and the co-occurrence of these features is maintained by homeostatic mechanisms that result from the exchange of genetic material through interbreeding and reproduction (Boyd, 1991, 288–289).

According to Darwin's theory of evolution by natural selection, phenotypic traits arise from variations within mechanisms occurring internally (e.g., genetic) and externally (e.g., environmental, ecological). Not every species member will be subjected to the same mechanisms; thus, different phenotypic traits arise at both the population level and within a given species population. Boyd (1999) notes that the HPC-Theory allows for phenotypic variations between species categorized in what he refers to as higher-level classifications. A higher-level classification categorizes species in a broader sense. For example, a lion and tiger fall into the species category ‘mammal’—despite their many differences, they share a cluster of properties and relations caused by mechanisms that reflect for example, the sharing of a common ancestor.

Like the essentialist theory, the HPC account of natural kinds

“retains a distinction between the observable properties of a kind and its explanatory core, but this core (i.e., the homeostatic mechanism) need not be either irreducible or intrinsic” (Pöyhönen, 2016, 150). An important feature of the HPC-Theory is that natural kinds are defined as the kinds of things that result from reliable and successful scientific practices. Boyd (1991) explains that because members of a particular natural kind category share a cluster of relevant properties or relations, they permit scientists to make legitimate generalizations and furthermore, aid in development of reliable intervention strategies that in turn, ground scientific explanations and predictions for the discovery of other natural kinds. In this sense, the mind-independent aspect of natural kinds is to some extent abandoned. According to Boyd (1991), a mind-dependent aspect plays a role in classification because scientists use methods that best serve their investigative aims. Boyd (1999, 1991) refers to this process that is central to the HPC-Theory, as “accommodation” in order to describe the connection between explanatory practices of kind use and the practice of kind category construction. The construction of natural kind categories is accommodated by the kinds of things that are epistemically useful since they feature in successful scientific practices. However, mind-independency is supposedly retained in the sense that natural kinds track causal structures in the world.

3.3. Mechanistic property clusters

The HPC-Theory has been adopted and slightly modified by a group of philosophers of science seeking to link psychiatric disorders to mechanisms in the brain, and as a way to construct a framework for psychiatric classification. Kendler et al. (2011) recognize psychiatric disorders as naturally occurring features of the causal structure of the brain and propose that they exist as ‘mechanistic property clusters’ (MPCs). They speculate that the symptoms that manifest as psychiatric disorders arise from groups of properties that co-occur, and whose co-occurrence is sustained by dysfunctional causal mechanisms (Kendler et al., 2011). Furthermore, they claim that these mechanisms span and interact across multiple levels—from the genetic to the environmental level—and produce detectable symptoms (Kendler et al., 2011). Causal interactions are also believed to occur between symptoms that arise at different levels, meaning some individuals who are diagnosed with depression, for example, and show signs of insomnia and guilt, will be predisposed to tiredness and suicidal ideation, respectively (Kendler et al., 2011). Specifically, they state that MPCs “are useful for prediction, explanation and control precisely because the kinds are sustained by causal mechanisms” (Kendler et al., 2011, p. 1147). Thus, they believe that characterizing psychiatric disorders as MPCs will enable the discovery of relevant properties that such disorders share, and in turn, will yield successful explanations of their etiology (see recent work by Khalidi, 2015; Borsboom et al., 2018; Neilsen & Ward, 2020; Shaffner & Tabb for alternative viewpoints).

3.4. Constructed kinds

In stark contrast to natural kinds realists, constructivists about natural kinds reject the idea that there are mind-independent divisions in the natural world out there to be discovered. They instead contend that natural kind categories are socially constructed—that is, they are human inventions. Proponents of constructivism typically defend three theses about natural kinds: (1) purported natural kinds do not reflect the natural structure of the world; (2) categories are constructed and shaped by social, cultural, political, historical, and economic factors; and (3) the construction of categories is not objective; rather, they are contingent on the ‘interest-driven’ classificatory practices relative to different scientific disciplines.⁷ Perhaps the most prominent

constructivist within the philosophical literature on natural kinds is Hacking (1990, 1991, 1992a, 1992b, 1995b, 1999a, 1999b, 2004, 2007). As Hacking (1999a, 33) eloquently puts it, “the world does not come quietly wrapped up in facts. Facts are the consequences of ways in which we represent the world.” Specifically, he contends that different systems of classification “are not determined by how the world is, but are convenient ways in which to represent it” (Hacking, 1999a, 33).

Hacking uses the concept of ‘human kinds’ to distinguish phenomena under study in human sciences like psychology and psychiatry from the kinds of phenomena under study in the natural sciences like physics and biology. The study of human behaviors, mental conditions, temperaments, societal groups, and so on, involves classification systems that group individuals into categories such as ‘alcoholic’, ‘hyperactive child’, and ‘refugee’ (Hacking, 1995b, 1999a). Hacking notes that these “human kind” categories share the same epistemic function as natural kind categories in that they involve the use of “classifications that could be used to formulate general truths about people; generalizations sufficiently strong that they seem like laws about people, their actions, or their sentiments” (Hacking, 1995a, 352). However, he argues that unlike natural kinds, human kinds do not track a stable natural order (Hacking, 1995a, 1995b, 1999a; Kendig, 2016a, 2016b). The reason, according to Hacking (1995a), is that human kinds are subject to what he dubs ‘looping-effects’—interactive causal relationships that occur between classified individuals and classificatory practices.

In the case of *human kinds*, when an individual is classified they may become aware of it, and “they can make tacit or even explicit choices, adapt or adopt ways of living so as to fit or get away from the very classification that may be applied to them” (Hacking, 1995a, 32). Depending on whether the classification is perceived negatively or positively by society, individuals may be motivated to change accordingly. Hacking (1995a, 1999a) contends that because people change their behaviors and attitudes over time, looping effects not only undermine the classification but also demonstrate that contrary to traditional models of natural kinds—wherein a single set or cluster of properties or relations are required—people simply do not conform to a single human kind category. Moreover, it is not only the individuals themselves who respond to being classified, people around them also respond and have an impact, too. Hacking (1995b, 103) argues that looping effects occur within the “larger matrix of institutions and practices surrounding the classification.” For example, in the case of hyperactive children, their behaviors may change in response to the actions, behaviors, and attitudes of their parents, peers, teachers, and doctors. In such cases, categories emerge and shift in response to social institutions and norms.

Advocates of looping effects use them to argue that mental disorders are not natural kinds (see Cooper, 2004 for an alternative view). Opponents charge constructivists with neglecting scientific evidence that some mental illnesses are attributable to underlying biological factors, not social convention. Khalidi (2013), for example, suggests that Hacking’s constructivist stance is misguided in positing a dichotomy between the natural sciences and the human sciences. Khalidi claims instead that natural kinds are common in the human sciences because the descriptive categories they use rely on the capacity of inductive practices to successfully track objective structures of reality (Khalidi, 2013; see also Dupré, 2004; Guala, 2010). Kendler et al. (2011, 1145) suggest classification in psychiatry “should seek common biological, psychological and social factors that warrant extrapolation across cultural and historical contexts.” Generally speaking, realists about natural kinds have not been swayed by Hacking’s arguments and still hold fast to the possibility of discovering natural kinds of *some form* in the special

(footnote continued)

and interest-relative pragmatism play some role in shaping scientific classification and thus in the identification of the kinds that areas of science discover. In psychiatry in particular, some scholars (Neilsen, 2020) contend that clinical diagnoses necessarily involve value judgments.

⁷ Many philosophers who advocate for natural kinds believe that convention

and human sciences.

From one vantage point, natural kinds and constructed kinds may be regarded as being at opposite ends of the kind continuum—either our categories pick out real mind-independent natural divisions in the causal structure of the world or all of the kinds picked out by human engineered concepts are mind-dependent. There is, however, a growing consensus among natural kinds realists that human interests, aims, values and practices play an ineliminable role in kind individuation in science—a role that given the satisfaction of certain constraints, need not compromise the reality or naturalness of those kinds (e.g., Boyd, 2000; Craver, 2009; Franklin-Hall, 2015; Kendig, 2016a; Khalidi, 2013; Weiskopf, 2020). This has prompted some philosophers to replace the designator “natural kinds” with, for example, “practical kinds” (e.g., Zachar, 2002, 2014, 2015), “investigative kinds” (e.g., Brigandt, 2003, 2012; Griffiths, 2004) and “categorical bottlenecks” (Franklin-Hall, 2015). We end this section by briefly considering some of these developments.

Advocates of essentialism and homeostatic or mechanistic property clusters insist that we should strive to build classifications grounded on features that are shared across category members, discovered by empirical investigation, and not constructed by fiat. Pragmatists, however, point out that scientific concepts often have practical goals associated with different individuals and organizations and just so long as these categories allow the fulfillment of these goals, they may be regarded as “natural” (e.g., Franklin-Hall, 2015). This “practical kinds” or instrumentalist approach is far less restrictive than other theories of natural kinds, because it allows a wider range of categories to be considered natural while easily ruling out arbitrary categories that can be shown to serve no practical aims (Franklin-Hall, 2015; Zachar, 2002). According to such accounts, classificatory practices proceed as scientists assess which categories best meet their practical goals, such as discovering and establishing the etiology of disorders, reliable diagnoses, and effective treatments (Kendler et al., 2011; Zachar, 2002, 2014, 2015). Categories are judged exclusively on the basis of their practical success, rather than their correspondence to a mind-independent reality. Peter Zachar, for example, argues that mental illness categories are best understood as constituted by an interaction “between what the world produces and what we find useful to notice” (Zachar, 2015, 289). Critics of the practical kinds view argue, however, that practical successes sometimes outweigh matters of fact when adopting classifications, and that shifts in interests and goals may prompt changes to classification systems in ways that compromise the naturalness of the kinds they pick out (citation).

The idea that scientific classification “reflects the immense variety of human interests,” is an indication to some philosophers that there is a plurality of ways to classify natural kinds (Dupré, 1981, p. 80). According to John Dupré’s (1993, 18) notion of *promiscuous realism* “there are countless legitimate, objectively grounded ways of classifying objects in the world.” Defenders of pluralism about natural kinds acknowledge that natural kinds are not only found within the natural sciences (Dupré, 1993; Khalidi, 2013). Pluralists also take into account the diversity of ways in which scientific disciplines develop constructs for investigative purposes (e.g., Dupré, 1993; Kendig, 2016a, 2016b; Khalidi, 2013; Sullivan, 2017a, 2017b). For example, schizophrenia is studied by geneticists, doctors, sociologists, anthropologists, philosophers, neuroscientists and so on—each of which analyse constructs of schizophrenia in different ways and for different purposes. Thus, it is argued that there are diverse ways of classifying kinds, and in some instances, there may be kinds that fit into more than one natural kind category (e.g., Kendig, 2016a, 2016b)—kinds that cross-cut each other or overlap (e.g., Dupré, 1993; Khalidi, 2013, 1998). As Khalidi contends, “if classification is always relative to certain interests, we would expect some [categories] to reorganize some of the same entities in different ways without displacing existing ones” (Khalidi, 1998, p. 42).

Philosophical understandings of the concept of natural kinds and debates about the usefulness of the very concept for understanding

scientific classification and conceptual change are still evolving as philosophers of science expand the focus of their inquiry to a number of diverse and interdisciplinary areas of science (see e.g., Bolker, 2013; Brigandt, 2003, 2010, 2012; Bursten, 2016; Godman, 2013; Kendig, 2016a, 2016b; Ludwig, 2017; Ludwig, 2018; Muszynski & Malaterre, 2020; Ruphy, 2010; Slater, 2015; Slater, 2013; Tabb, 2019; Tsou, 2013; Zachar, 2000). Natural kinds realists have expressed optimism that just so long as the aims of classification in a given scientific domain are broadly epistemic, natural kinds in some form (e.g., HPC, MPC) will be in the offing (e.g., Boyd, 2019; Kendler et al., 2011; Khalidi, 2013), and there may be different epistemically admirable ways of conceptually carving up the world that cross-cut each other (e.g., Khalidi, 2013). However, debates about the status of kinds in psychology, neuroscience and psychiatry are on-going. As we explain in Section 4, setting aside the possibility that the kinds of things under study in these areas of science (e.g., mental illnesses, cognitive functioning) may be subject to looping effects (Hacking, 1995a, 1995b) or the reactivity (e.g. Harris & Lahey, 1982) of experimental subjects (e.g., humans, non-human primates, rodents), the “kinding practices” (e.g., Kendig, 2016b) of investigators also have important implications for classification and the status of kinds in these areas of science (see also e.g., Chang, 2004, Chang, 2017 on natural kinds and “epistemic iteration”).

4. Scientific practice, coordinated pluralism and coordinated kinds

If we synthesize the ideas of those philosophers we have considered in the previous two sections, classification in science may be roughly described as proceeding in stages. First, conventional “folk” categories are abandoned for descriptive systems of classification that specify criteria of application for scientific terms based on objectively verifiable features of objects. These classification systems do not track “natural kinds”, but they are regarded as important starting points for causal discovery insofar as they ensure scientists hailing from different theoretical backgrounds, who take themselves to be interested in roughly the same domain of phenomena, share a set of conceptual categories in common. These consensus-based operationalized concepts may then serve as a basis for integrating results from multiple different domains of scientific inquiry. Then, descriptive classification systems gradually are replaced or updated in light of causal discovery with an eye towards categories that facilitate explanation, prediction, control and understanding. Some philosophers believe that the resulting categories track “natural kinds” and have put forward different understandings of this concept, conceding that the kinds that are discovered are partly due to our “workmanship” in accommodating our concepts to the world (e.g., Boyd, 2000).

This account of how classification proceeds in science is highly idealized; if natural kinds are ultimately the goal of science, investigations into the nature of *scientific practice* reveal that classificatory practices and kind discovery in science do not follow this logical trajectory (e.g., Andersen, 2010; Bloch-Mullins, 2020a, 2020b, Chang, 2004, 2012, 2011; Hacking, 1992b; Feest & Steinle, 2012; Kendig, 2016a, 2016b; Nersessian, 2008). Numerous philosophers of science and scientists have noted that exploratory and hypothesis-driven research often proceed in science without a firm grasp of what the objects of inquiry are nor how the kinds of things under study fit into some broader taxonomy of kinds (e.g., Anderson, 2015; Brigandt, 2003, 2020; Chang, 2004, 2012; Colaço, 2020; Feest, 2005, 2010, 2011, 2012, 2017; Feest & Steinle, 2012; Griffiths, 1997, 2004; Haueis & Slaby, 2017; Kendig, 2016a, 2016b; Muszynski & Malaterre 2020; Rheinberger, 1997; Sullivan, 2009, 2017b, Sullivan, 2010; Sullivan, 2019; Sullivan, 2020).

Especially in interdisciplinary areas of science, such as evolutionary development biology, cognitive neuroscience, criminology, clinical psychiatry, ethnobiology neurobiology to name only a handful, we encounter a host of practices that do not appear to be aligned with the

natural kinds ideal. Consider an example from translational cognitive neuroscience, which aims to translate findings about the mechanisms of impairments in cognitive functioning from animal models of disease to humans suffering from these impairments. It is widely recognized that many mental illnesses, neurodegenerative diseases and brain disorders (e.g., concussion) are accompanied by impairments in cognitive functions like memory, attention, and decision-making. Patients with these disorders may have difficulty dividing, switching or maintaining attention, remembering objects, persons and locations or adjusting their behavior when appropriate. In the first two decades of the 21st century, two major initiatives were established in the United States with an eye towards propelling forward the discovery of novel therapeutics for these cognitive impairments. Although the *Research Domain Criteria Project* (e.g., Cuthbert & Kozak, 2013; Morris and Cuthbert, 2012; Sanislow et al., 2010; Insel et al., 2010) is the better known of the two, our focus here is on the *Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia* (CNTRICS) initiative (e.g., Carter & Barch, 2007; Moore et al., 2013). This initiative variously brought together preclinical translational behavioral neuroscientists, clinical researchers, cognitive neuroscientists working with humans and/or animal models, systems neuroscientists and members of the pharmaceutical industry to (1) identify the physiological, genetic and developmental mechanisms that underlie cognitive impairments (e.g., deficits in memory, attention) in Schizophrenia and to (2) locate effective therapeutics for treating such impairments. CNTRICS was aimed in particular at developing testing methods (behavioral assays) for assessing cognition and cognitive impairments and their mechanisms in animal models of disease (rodents) that would facilitate the translation of pre-clinical findings to human patients in the form of effective therapeutic interventions (e.g., Hvoslef-Eide, Nilsson, Saksida, & Bussey, 2015; Moore et al., 2013).

Across a number of large-scale and working group meetings, scientists Deanna Barch and Cameron Carter, who spearheaded the CNTRICS initiative, identified in collaboration with other investigators, a number of hurdles to their explanatory, predictive and therapeutic goals. In order to illustrate the nature of these hurdles, consider an example. One cognitive function that is widely believed to be impaired in Schizophrenia and other mental illnesses is *cognitive control* (McTeague, Goodkind, & Etkin, 2016, 2017). This concept is variously defined as “the top-down modulation of cognitive processes based on higher-order representations such as goals or plans” (<https://www.cognitiveatlas.org>), “the process by which goals or plans influence behaviour” (<https://www.nature.com/subjects/cognitive-control>), and “a system that modulates the operation of other cognitive and emotional systems, in the service of goal-directed behavior, when prepotent modes of responding are not adequate to meet the demands of the current context” (<https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc/constructs/cognitive-control.shtml>). These general definitions of the construct are not identical and allow for the possibility of a wide array of phenomena to be classified as instances of cognitive control. A number of different tasks, including the *Wisconsin Card Sorting Task* (WCST), the digit span task, and the Porteus maze test, to name only a few, may be used to study cognitive control in humans (see www.cognitiveatlas.org). Task demands differ across these tasks, leaving open the possibility that the tasks measure different kinds of cognitive processes. The WCST, for example, has been used to measure other cognitive functions including task switching and set shifting (<https://www.cognitiveatlas.org>). This is suggestive that different investigators have different intuitions about what cognitive function the WCST actually measures. On one version of the WCST, a subject must sort a set of cards containing different stimuli into separate piles without knowing precisely which stimulus dimensions are to be used to sort the cards. The subject is required to rely on input from the task administrator, who only indicates whether a given choice of sorting is correct or incorrect. The subject also may be tested under conditions in which they have learned one set of rules, but then the

rules are changed and they have to learn a new set of rules also exclusively on the basis of administrator feedback (<https://www.cognitiveatlas.org>). The task has a number of different parameters, and investigators who use it are at liberty to vary stimulus parameters and other features of the experimental protocol as they deem relevant to their research goals.

The aforementioned features are not unique to the concept of cognitive control or the WCST (see e.g., Cuthbert & Kozak, 2013; Hommel et al., 2019; Irvine, 2012; Poldrack, 2010; Poldrack et al., 2011; Stinson, 2009; Sullivan, 2009, 2017a, 2017b). In fact, researchers involved in the CNTRICS initiative identified a number of important obstacles to translational research that are generally applicable across cognitive concepts and cognitive assessment tools used in contemporary neuroscience. To summarize, it is not uncommon for investigators hailing from different theoretical traditions to use the same concepts to individuate kinds of phenomena, but to mean different things by these concepts. In other words, concepts in neuroscientific areas that study human and non-human animal cognition lack *interrater reliability*, or consistent criteria of application across investigators (see Haslam, 2013; Hempel, 1965). This is a barrier to discovery insofar as it promotes misunderstandings among researchers who may otherwise regard themselves as engaged in a common explanatory project. There is also no widespread consensus about which tests or tasks should be used to assess which cognitive functions and investigators have the freedom to modify various aspects of the protocols used in conjunction with tests and tasks as they deem appropriate to their explanatory goals (e.g., Poldrack, 2010; Poldrack and Yarkoni, 2016; Sullivan, 2009). A related issue is that cognitive assessment tools used to individuate kinds of cognitive functions and intervention tools used to probe for causes are often not standardized across investigators. Given that even subtle differences in experimental techniques may result in large differences in the data such techniques produce, it is an open question whether results from different experiments purportedly investigating a single cognitive process may be integrated or synthesized into the same explanation of the same phenomenon (e.g., Sullivan, 2009).

What implications might the aforementioned features of scientific practice have for the discovery of so-called natural kinds in translational areas of cognitive neuroscience (as well as in other areas of science where we encounter similar issues)? One possibility is that conceptual and methodological pluralism is so rampant and unbridled that investigators are not progressing towards anything like the natural kinds ideal (e.g., Sullivan, 2016b). However, researchers involved in the CNTRICS initiative do not regard widespread conceptual and methodological pluralism as an adequate stopping point for translational research. Rather, as we explain in the remainder of this section, they may be understood as advocating for a set of consensus-based strategies and epistemic benchmarks for translational research that may at best be aligned with a property clusters account of the natural kinds ideal (just so long as that account is supplemented with descriptive details about the nature of scientific practice).

In order to overcome conceptual and methodological hurdles to the success of translational research, CNTRICS researchers sought to reach consensus on (a) a set of constructs designating cognitive functions and sub-functions and (b) a set of tasks that could be used to investigate these functions in humans and animal models (e.g., Barch et al., 2009; Carter & Barch, 2007; Barch & Carter, 2008; Carter, Kerns, & Cohen, 2009; Carter, Barch, & Buchanan, 2008; Moore et al., 2013). Two selection criteria for functions and tasks included *construct validity* and *neurocognitive validity* (e.g., Hvoslef-Eide et al., 2015). Specifically, across the contexts of in-person meetings and web-based surveys, CNTRICS participants were asked to identify “constructs that have prominence in the field of cognitive neuroscience and substantial promise for delineating elementary cognitive processes” (i.e., construct validity), which “may be more closely connected to neural systems” (i.e., *neurocognitive validity*) (Carter et al., 2008, 5). Yet, CNTRICS participants were open to the possibility that the constructs and definitions

they selected “were only one way of ‘carving nature at its joints’” (Carter et al., 2008, 9) and that other cognitive constructs were possible.

From the perspective of CNTRICS researchers, satisfying the criteria of construct validity and neurocognitive validity with respect to conceptual categories designating cognitive functions is necessary but not sufficient for the success of translational research. Researchers also must agree on a broader set of *epistemic benchmarks* that their cognitive assessment, intervention and data analysis tools must meet and a set of best practices for translational research. For example, an additional hurdle to translating results from rodent to human cognitive assessment studies and integrating findings that apply across species, is the mismatch between the tools used to assess cognition in rodents and humans (e.g., Bussey et al., 1994; Bussey et al., 2012). Many researchers involved in the CNTRICS initiative agree that the more similar the tasks for assessing cognition are across species, i.e., *face validity*, the more likely it is that the same cognitive functions and neural circuits will be involved. Yet even if the tasks used to probe cognitive functions in different species look similar, this does not guarantee they are suitable for individuating cognitive capacities (construct validity (e.g., Cronbach & Meehl, 1955; Slaney, 2017; Shadish et al., 2002; Sullivan, 2019) and identifying neural circuits that mediate task performance (*neurocognitive validity*, e.g., Hvoslef-Eide et al., 2015). CNTRICS participants reached some consensus that construct validity and neurocognitive validity with respect to cognitive assessment tools for use in humans and rodents are essential (e.g., Carter & Barch, 2007). Achieving these types of validity is, however, an iterative process, as findings from exploratory and hypothesis-driven experiments prompt task refinement and/or construct revision (e.g., NIMH, 2016; Shadish et al., 2002; Sullivan, 2016a; Hvoslef-Eide et al., 2015).

The success of translational research is also thought to require establishing that task performance *across species* recruits the same neural circuits. This is referred to as *translational neurocognitive validity* (e.g., Hvoslef-Eide et al., 2015), which is fundamental for predicting whether functional outcomes of treatment interventions observed in one species will be similar in another. Ideally, translational researchers want to be able to predict that a treatment that works in rescuing a cognitive deficit in a rodent model of disease will be effective in rescuing that deficit in human clinical populations. Cognitive assessment tools should also be sensitive to dose-dependent effects of drugs on cognitive abilities, so as to increase the likelihood that a drug that ameliorates cognitive impairments in a rodent model has a greater chance of having a positive impact in the human case. Achieving translational neurocognitive validity also is an iterative process, as it requires investigators working in human and non-human animal cognition to move back and forth to refine human and rodent tasks in order to facilitate translation while ensuring that other dimensions of validity are maintained.

These varied dimensions of validity are emphasized in the CNTRICS literature. They are all intimately related, however, to another important desiderata for experiments: *reliability*. An experiment is ideally supposed to leave an investigator epistemically better off than she would have been had she never run that experiment. Improving one's epistemic situation by means of an experiment requires that the experiment be *reliable* or capable of producing data requisite to adjudicate among competing hypotheses about a cognitive capacity of interest and its mechanisms or the efficacy of a drug on improving an impairment in that capacity. Every cognitive assessment task or tool is one component of a broader experimental protocol or set of standardized operating procedures (SOP) that specifies the steps that must followed from the beginning to the end of each experiment in which that task or tool is used. The specified steps are ideally supposed to rule out the possibility of confounds, experimenter error and artifacts (see e.g., Mayo, 1996; Schikore, 2019; Sullivan, 2018; Sullivan, 2020). Experiments may be more or less reliable; increasing the reliability of an experiment is contingent upon an investigator actively thinking about and probing for potential kinds of confounds that may occur during the processes of

data collection and analysis. We have argued previously that adequate probing for errors requires “perspectival pluralism” (e.g., Giere, 2010; Sullivan, 2014, 2018; Wimsatt, 2007) and methodological pluralism (e.g., Sullivan, 2018) as input from investigators with different theoretical backgrounds and diverse technical expertise is used as a basis for itemizing potential confounds and errors to which experiments may be subject (e.g., Sullivan, 2014, 2018). A lack of reliability in exploratory experiments may jeopardize knowledge production and discovery; in hypothesis-driven experiments it may negatively impact the ability to adjudicate among competing hypotheses about the effects under study in the laboratory, and ultimately jeopardize the ability to generalize results beyond that context.

Additionally, experimental results must be *reproducible*. Many labs must run the same experiments in order to determine whether the effects observed in a single lab are real or idiosyncratic to a given experimental context. Researchers must also have good grounds for thinking that the same phenomenon is being investigated when different research groups use the same cognitive tasks. For all these conditions to be possible, CNTRICS researchers agreed that the cognitive assessment tools used in humans and/or rodents, including the operating procedures/protocols used in conjunction with them, must be standardized across research groups (See also *International Brain Laboratory*, 2017).

Meeting the aforementioned benchmarks is not possible in a single laboratory or research study; rather, it requires extensive collaboration and coordination of research practices within and across research groups working at the same and different levels of analysis and with different species. Such groups must include investigators having diverse theoretical backgrounds and technical expertise. A significant amount of resources and time are also required to gradually hit these benchmarks and to reach the point at which discovery of novel therapeutic interventions is possible. We refer to such collaboration as “coordinated pluralism” (Sullivan, 2017a; Mattu, 2020) and believe that just so long as these kinds of epistemic benchmarks are sought after in science, what may be discovered in conjunction with such practices may best be understood as “coordinated kinds”. We think the concept of “coordination”, in contrast to “accommodation” (e.g., Boyd, 2000), better emphasizes the kind of intensive collaborative work required on the part of scientists to progress towards something like the natural kinds ideal. To the extent that coordinated pluralism exists in science, the coordinated kinds associated with it will be more constrained groupings maintained by subsets of practitioners who aim to align their practices in specific ways to meet their explanatory and/or predictive goals (see Ankeny & Leonelli, 2015, 2016 on “scientific repertoires” and Kutschenko, 2011 on “epistemic hubs”). We believe that while conceptual and methodological pluralism are widespread in science, coordinated pluralism, in contrast, is quite rare.

5. Conclusion

In this paper, we provided an historical overview of the principles that philosophers have claimed have and ought to guide the development of classification systems in science. We went on to illustrate, through a survey of relevant philosophical literature, that philosophers of science have been historically interested in concept of natural kinds. However, given differences in the kinds studied in different areas of science, a number of different understandings of the concept of natural kind have been put forward and debates about the status of kinds in many areas of science are still on-going. We also considered a number of different understandings of kinds that have been developed to accommodate the unstable and more complex causal structure of kinds under study in “special sciences” like biology, psychology and psychiatry. In the final section of the paper, we emphasized that success in achieving “the natural kinds ideal” in some areas of science appears to be contingent on the conceptual and methodological practices of investigators and we described how the strategy of “coordinated

pluralism” might at best facilitate the discovery of “coordinated kinds” in science, which may be as close as some areas of science get to “natural” kinds.

Acknowledgements

The authors would like to thank two anonymous referees for very helpful comments on an earlier version of this paper. This project was funded by a Strategic Support for Tri-Council Success Grant from Western University, London, Ontario, Canada awarded to Jacqueline Sullivan.

References

- Ackrill, J. L. (1963). *Aristotle: Categories and De Interpretatione*. Oxford: Clarendon Press.
- Andersen, H. (2010). Joint acceptance and scientific change: A case study. *Episteme*, 7, 248–265.
- Anderson, M. (2015). Mining the brain for a new taxonomy of the mind. *Philosophy Compass*, 10(1), 68–77.
- Ankeny, R., & Leonelli, S. (2015). Repertoires: How to transform a project into a research community. *Bioscience*, 65(7), 701–708.
- Ankeny, R., & Leonelli, S. (2016). Repertoires: A post-Kuhnian perspective on scientific change and collaborative research. *Studies in History and Philosophy of Science*, 60, 18–28.
- Bacon, Francis (2000). In Lisa Jardine, & Michael Silverthorne (Eds.). *The New Organon*. Cambridge, UK: Cambridge University Press.
- Barch, D., & Carter, C. (2008). Measurement issues in the use of cognitive neuroscience tasks in drug development for impaired cognition in schizophrenia: A report of the second consensus building conference of the CNTRICS initiative. *Schizophrenia Bulletin*, 34(4), 613–618.
- Barch, D., Carter, C., Arnsten, A., et al. (2009). Selecting paradigms from cognitive neuroscience for translation into use in clinical trials: Proceedings of the third CNTRICS meeting. *Schizophrenia Bulletin*, 35(1), 109–114.
- Bird, A. (2010). Essences and natural kinds. In H. Beebe, & N. Sabbarton-Leary (Eds.). *The semantics and metaphysics of natural kinds* (pp. 125–136). Routledge.
- Bird, A., & Tobin, E. (2018). Natural kinds. In E. N. Zalta (Ed.). *The Stanford encyclopedia of philosophy* <https://plato.stanford.edu/archives/spr2018/entries/natural-kinds>.
- Bloch-Mullins, C. (2020a). Concepts as forward-looking: How taxonomic structure facilitates conceptual development. *Journal of the Philosophy of History*, 1–27.
- Bloch-Mullins, C. (2020b). Similarity reimagined (with implications for a theory of concepts). *Theoria*. <https://doi.org/10.1111/theo.12197>.
- Mattu, J. (2020). Integrating Taxonomies of Emotion: ‘Coordinated Pluralism’ in Emotion Research.
- Bolker, J. (2013). The use of natural kinds in evolutionary developmental biology. *Biological Theory*, 7, 121–129.
- Borsboom, D., Cramer, A., & Kalis, A. (2018). Brain disorders? Not really... Why network structures block reductionism in psychopathology research. *Behavioral and Brain Sciences*, 1–54. <https://doi.org/10.1017/S0140525X17002266>.
- Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies*, 61, 127–148.
- Boyd, R. (1999). Homeostasis, species, and higher taxa. In R. A. Wilson (Ed.). *Species: New interdisciplinary essays* (pp. 141–186). Cambridge, MA: MIT Press.
- Boyd, R. (2000). Kinds as the workmanship of men: Realism, constructivism, and natural kinds. In J. Nida-Rümelin (Ed.). *Rationalität, Realismus, Revision* (pp. 52–89). Berlin: De Gruyter.
- Boyd, R. (2019). Rethinking natural kinds, reference and truth: Towards more correspondence with reality, not less. *Synthese*. <https://doi.org/10.1007/s11229-019-02138-4>.
- Bridgman, P. (1927). *The logic of modern physics*. New York: Macmillan.
- Bridgman, P. (1938). Operational analysis. *Philosophy of Science*, 5, 114–131.
- Bridgman, P. (1952). *The nature of some of our physical concepts*. New York: Philosophical Library ((Reprinted from the British Journal for the Philosophy of Science, 1 (1950): 257–272; 2 (1951): 25–44) (1952)).
- Brigandt, I. (2003). Species pluralism does not imply species eliminativism. *Philosophy of Science*, 70, 1305–1316.
- Brigandt, I. (2010). The epistemic goal of a concept: Accounting for the rationality of semantic change and variation. *Synthese*, 177, 19–40.
- Brigandt, I. (2012). The dynamics of scientific concepts: The relevance of epistemic aims and values. In U. Feest, & F. Steinle (Eds.). *Scientific concepts and investigative practice* (pp. 75–103). Berlin: Walter de Gruyter.
- Brigandt, I. (2020). How to philosophically tackle kinds without talking about natural kinds. *Canadian Journal of Philosophy*. <https://doi.org/10.1017/can.2020.29>.
- Bursten, J. (2016). Smaller than a breadbox. *The British Journal for the Philosophy of Science*. <https://doi.org/10.1093/bjps/axw022>.
- Bussey, T., Muir, J., & Robbins, T. (1994). A novel automated touchscreen procedure for assessing learning in the rat using computer graphic stimuli. *Neuroscience Research Communications*, 15(2), 103–110.
- Bussey, T. J., Holmes, A., Lyon, L., et al. (2012). New translational assays for preclinical modelling of cognition in schizophrenia: The touchscreen testing method for mice and rats. *Neuropharmacology*, 62(3), 1191–1203.
- Carter, C., Barch, D., Buchanan, E., et al. (2008). Identifying cognitive mechanisms targeted for treatment development in schizophrenia: an overview of the first meeting of the Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia Initiative. *Biological Psychiatry*, 64(1), 4–10.
- Carter, C., Kerns, J., & Cohen, J. (2009). Cognitive neuroscience: bridging thinking and feeling to the brain, and its implications for psychiatry. *Neurobiology of Mental Illness*, 3rd edition, D. Charney and E. Nestler (pp. 168–178). Oxford, UK: Oxford University Press.
- Carter, C. S., & Barch, D. M. (2007). Cognitive neuroscience-based approaches to measuring and improving treatment effects on cognition in schizophrenia: The CNTRICS initiative. *Schizophrenia Bulletin*, 33(5), 1131–1137.
- Chang, H. (2004). *Inventing temperature: Measurement and scientific progress*. New York: Oxford University Press.
- Chang, H. (2012). Is water H₂O? Evidence, realism and pluralism. *Boston studies in the philosophy and history of science*. Cambridge, UK: Springer.
- Chang, H. (2019). Operationalism. *Stanford encyclopedia of philosophy*.
- Chang, Hasok (2017). Epistemic iteration and natural kinds: Realism and pluralism in taxonomy. In K. S. Kendler, & J. Parnas (Eds.). *Philosophical Issues in Psychiatry IV: Psychiatric Nosology* (pp. 229–245). Oxford, UK: Oxford University Press.
- Churchland, P. (1981). Eliminative materialism and the propositional attitudes. *The Journal of Philosophy*, 78(2), 67–90.
- Colaço, D. (2020). Recharacterizing scientific phenomena. *European Journal for Philosophy of Science*, 10(2), 1–19.
- Cooper, R. (2004). Why hacking is wrong about human kinds. *British Journal for the Philosophy of Science*, 55, 73–85.
- Craver, C. (2009). Mechanisms and natural kinds. *Philosophical Psychology*, 22(5), 575–594.
- Cronbach, L., & Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302.
- Cuthbert, B. N., & Kozak, M. J. (2013). Constructing constructs for psychopathology: The NIMH research domain criteria. *Journal of Abnormal Psychology*, 122(3), 928–937.
- Devitt, M. (2008). Resurrecting biological essentialism. *Philosophy of Science*, 75, 344–382.
- Dupré, J. (1981). Natural kinds and biological taxa. *Philosophical Review*, 90(1), 66–90.
- Dupré, J. (1993). *The disorder of things: Metaphysical foundations of the disunity of science*. Cambridge, MA: Harvard University Press.
- Dupré, J. (2004). Human kinds and biological kinds: Some similarities and differences. *Philosophy of Science*, 71(5), 892–900.
- Ellis, B. (2001). *Scientific essentialism*. Cambridge University Press.
- Ellis, B. (2008). Essentialism and natural kinds. In M. Curd, & S. Psillos (Eds.). *The Routledge companion to philosophy of science* (pp. 139–149). Routledge.
- Ereshfsky, M. (2004). Bridging the gap between human kinds and biological kinds. *Philosophy of Science*, 71(5), 912–921.
- Feest, U. (2005). Operationism in psychology: What the debate is about, what the debate should be about. *Journal of the History of the Behavioral Sciences*, 41(2), 131–149.
- Feest, U. (2010). Concepts as tools in the experimental generation of knowledge in cognitive neuropsychology. *Spontaneous Generations: A Journal for the History and Philosophy of Science*, 4(1), 173–190.
- Feest, U. (2011). Remembering (short-term) memory: Oscillations of an epistemic thing. *Erkenntnis*, 75(3), 391–411.
- Feest, U. (2012). Exploratory experiments, concept formation and theory construction in psychology. *Scientific concepts and investigative practice* (pp. 167–189). Berlin: De Gruyter.
- Feest, U. (2017). Phenomena and objects of research in the cognitive and behavioral sciences. *Philosophy of Science*, 84(5), 1165–1176.
- Feest, U., & Steinle, F. (2012). *Scientific concepts and investigative practice*. Berlin: De Gruyter.
- Franklin-Hall, L. (2015). Natural kinds as categorical bottlenecks. *Philosophical Studies*, 172(4), 925–948.
- Giere, R. (2010). *Scientific perspectivism*. Chicago: University of Chicago Press.
- Godman, M. (2013). Psychiatric disorders qua natural kinds. *Biological Theory*, 144–152.
- Griffiths, P. (1997). *What emotions really are: The problem of psychological categories*. Chicago: University of Chicago Press.
- Griffiths, P. (2004). Emotions as natural and normative kinds. *Philosophy of Science*, 71(5), 901–911.
- Grube, G. M. A. (1974). *Plato's republic*. Indianapolis: Hackett Publishing Co.
- Guala, F. (2010). Infallibilism and human kinds. *Philosophy of the Social Sciences*, 40(2), 244–264.
- Hacking, I. (1990). Natural kinds. In R. Barrett, & R. Gibson (Eds.). *Perspectives on quine* (pp. 129–141). Oxford: Blackwell.
- Hacking, I. (1991). A tradition of natural kinds. *Philosophical Studies*, 61(2), 109–126.
- Hacking, I. (1992a). World-making by kind-making: Child abuse for example. In M. Douglas, & D. L. Hull (Eds.). *How classification works: Nelson Goodman among the social sciences* (pp. 180–238). Edinburgh: Edinburgh University Press.
- Hacking, I. (1992b). The self-vindication of the laboratory sciences. In A. Pickering (Ed.). *Science as practice and culture*. University of Chicago Press.
- Hacking, I. (1995a). The looping effects of human kinds. In D. Sperber, D. Premack, & A. J. Premack (Eds.). *Causal cognition: A multidisciplinary debate* (pp. 351–383). Oxford: Clarendon Press.
- Hacking, I. (1995b). *Rewriting the soul: Multiple personality and the sciences of memory*. Princeton, NJ: Princeton University Press.
- Hacking, I. (1999a). *The social construction of what*. Cambridge, MA: Harvard University Press.
- Hacking, I. (1999b). Making up people. In M. Biagioli (Ed.). *The science studies reader* (pp. 161–171). New York: Routledge.
- Hacking, I. (2004). Between Michel Foucault and Erving Goffman: Between discourse in the abstract and face-to-face interaction. *Economy and Society*, 33(3), 277–302.

- Hacking, I. (2007). Natural kinds: Rosy dawn, scholastic twilight. *Royal Institute of Philosophy Supplement*, 61, 203–239.
- Harris, F., & Lahey, B. (1982). Subject reactivity in direct observational assessment: A review and critical analysis. *Clinical Psychology Review*, 2(4), 523–538.
- Haslam, N. (2013). Reliability, validity, and the mixed blessings of operationalism. In K. W. M. Fulford, M. Davies, R. Gipps, G. Graham, J. Sadler, G. Tanghellini, & T. Thornton (Eds.). *The Oxford handbook of philosophy and psychiatry*. Oxford: Oxford University Press.
- Hauens, P., & Slaby, J. (2017). Connectomes as constitutively epistemic objects: Critical perspectives on modeling in current neuroanatomy. *Progress in Brain Research*, 233, 149–177.
- Hempel, C. G. (1965). *Fundamentals of Taxonomy. Aspects of scientific explanation: And other essays in the philosophy of science*. New York, NY: The Free Press.
- Herschel, J. (1830 [2009]). *A preliminary discourse on the study of natural philosophy*. Cambridge: Cambridge University Press.
- Hommel, B., Chapman, C., Cisek, P., Neyedli, H., Song, J., & Welsh, T. (2019). No one knows what attention is. *Attention, Perception & Psychophysics*, 81(7), 2288–2303.
- Hull, D. L. (1965). The effect of essentialism on taxonomy: Two thousand years of stasis. *British Journal for the Philosophy of Science*, 15(60), 314–326.
- Hvoslef-Eide, M., et al. (2015). Cognitive translation using the rodent touchscreen testing approach. *Current Topics in Behavioral Neurosciences*, 28, 423–447.
- Insel, T., Cuthbert, B., Garvey, M., et al. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *The American Journal of Psychiatry*, 167(7), 748–751.
- International Brain Laboratory (2017). International Brain Laboratory. An international laboratory for systems and computational neuroscience. *Neuron*, 96(6), 1213–1218.
- Irvine, E. (2012). *Consciousness as a scientific concept*. Dordrecht: Springer.
- Kendig, C. (2016a). *Natural kinds and classification in scientific practice*. London: Routledge.
- Kendig, C. (2016b). Activities of kinding in scientific practice. *Natural kinds and classification in scientific practice*. London: Routledge.
- Kendler, K., Zachar, P., & Craver, C. F. (2011). What kinds of things are psychiatric disorders. *Psychological Medicine*, 41, 1143–1150.
- Khalidi, M. A. (1998). Natural kinds and crosscutting categories. *Journal of Philosophy*, 95(1), 33–50.
- Khalidi, M. A. (2013). *Natural categories and human kinds: Classification in the natural and social sciences*. Cambridge: Cambridge University Press.
- Khalidi, M. A. (2015). Natural kinds as nodes in causal networks. *Synthese*. <https://doi.org/10.1007/s11229-015-0841-y>.
- Kincaid, H., & Sullivan, J. (2014). *Classifying psychopathology: Mental kinds and natural kinds*. Cambridge, MA: MIT Press.
- Kutschenko, L. (2011). How to make sense of broadly applied medical classification systems: Introducing epistemic hubs. *History and Philosophy of the Life Sciences*, 33(4), 583–601.
- LaPorte, J. (2004). *Natural kinds and conceptual change*. Cambridge: Cambridge University Press.
- Locke, J. (1689 [1975]). In P. H. Niddich (Ed.). *An essay concerning human understanding*. Oxford: Clarendon Press (1975).
- Ludwig, D. (2017). Indigenous and scientific kinds. *The British Journal for the Philosophy of Science*, 68, 187–212.
- Ludwig, D. (2018). Letting go of “natural kind” towards a multidimensional framework of non-arbitrary classification. *Philosophy of Science*, 85(1), 31–52.
- Magnus, P. D. (2012). *Scientific enquiry and natural kinds: From planets to mallards*. Basingstoke: Palgrave Macmillan.
- Magnus, P. D. (2014). No grist for Mill on Natural Kinds. *Journal for the History of Analytical Philosophy*, 2(4), 1–15.
- Mayo, D. (1996). *Error and the growth of experimental knowledge*. Chicago: University of Chicago Press.
- McTeague, L., Goodkind, M., & Etkin, A. (2016). Transdiagnostic impairment of cognitive control in mental illness. *Journal of Psychiatric Research*, 83, 37–46.
- McTeague, L., Goodkind, M., & Etkin, A. (2017). Transdiagnostic impairment in cognitive control neurocircuits: Behaviour, structure, and function. In T. Egner (Ed.). *The Wiley handbook of cognitive control* (pp. 539–555). Wiley-Blackwell.
- Mellor, D. (1977). Natural kinds. *The British Journal for the Philosophy of Science*, 28, 299–312.
- Mill, J. S. (1874). *A system of logic* (8th ed.). New York: Harper & Brothers.
- Moore, H., Geyer, M. A., Carter, C. S., & Barch, D. M. (2013). Harnessing cognitive neuroscience to develop new treatments for improving cognition in schizophrenia: CNTRICS selected cognitive paradigms for animal models. *Neuroscience and Biobehavioral Reviews*, 9(Pt B), 2087–2091.
- Morris, S., & Cuthbert, B. (2012). Research domain criteria: Cognitive systems, neural circuits, and dimensions of behavior. *Dialogues in Clinical Neuroscience*, 14, 29–37.
- Muszynski, E., & Malaterre, C. (2020). Best Behavior: A proposal for a non-binary conceptualization of behavior in biology. Article 101222 *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 79, 1–10.
- National Institute of Mental Health, National Advisory Mental Health Council Workgroup on Tasks and Measures for Research Domain Criteria (RDoC) (2016). **Behavioral assessment of methods for RDoC constructs**. https://www.nimh.nih.gov/about/advisory-boards-and-groups/namhc/reports/rdoc_council_workgroup_report_153440.pdf.
- Neilsen, K. (2020). *What is a mental disorder? Developing and embodied, embedded and enactive psychopathology* (Doctoral Dissertation) Wellington NZ: Victoria University of Wellington.
- Neilsen, K., & Ward, T. (2020). Phenomena complexes as targets of explanation in psychopathology: The relational analysis of phenomena approach. *Theory and Psychology*, 30(2), 164–185.
- Nersessian, N. (2008). *Creating scientific concepts*. Cambridge, MA: MIT Press.
- Poldrack, R. (2010). Subtraction and beyond: The logic of experimental designs for neuroimaging. In S. Hanson, & M. Bunzl (Eds.). *Foundational issues in human brain mapping* (pp. 147–159). Cambridge, MA: MIT Press.
- Poldrack, R., Kittur, A., Kalar, D., Miller, E., Seppa, C., Gil, Y., ... Bilder, R. (2011). The cognitive atlas: Toward a knowledge foundation for cognitive neuroscience. *Frontiers in Neuroinformatics*, 5, 17.
- Poldrack, R., & Yarkoni, T. (2016). From brain maps to cognitive ontologies: Informatics and the search for mental structure. *Annual Review of Psychology*, 67, 587–612.
- Pöyhönen, S. (2016). Memory as a cognitive kind. In C. Kendig (Ed.). *Natural kinds and classification in scientific practice* (pp. 145–156). London: Routledge.
- Quinn, A. (2017). Whewell on classification and consilience. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 64, 65–74.
- Rheinberger, H.-J. (1997). *Towards a history of epistemic things: Synthesizing proteins in the test tube*. Stanford, CA: Stanford University Press.
- Ruphy, S. (2010). Are stellar kinds natural kinds. *Philosophy of Science*, 77, 1109–1120.
- Sanislow, C., Pine, D., Quinn, K., Kozak, M., Garvey, M., Heissen, R., ... Cuthbert, B. (2010). Developing constructs for psychopathology research: Research domain criteria. *Journal of Abnormal Psychology*, 119(4), 631–639.
- Schikore, J. (2019). The structure and function of experimental control in the life sciences. *Philosophy of Science*, 86(2), 203–218.
- Shadish, W., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Slaney, K. (2017). *Validating psychological constructs, historical, philosophical, and practical dimensions*. London: Palgrave MacMillan.
- Slater, M. (2013). Cell types as natural kinds. *Biological Theory*, 7, 170–179.
- Slater, M. (2015). Natural kindness. *The British Journal for the Philosophy of Science*, 66, 375–411.
- Stinson, C. (2009). Searching for the source of executive attention. *Psyche*, 15(1), 137–154.
- Sullivan, J. (2009). The multiplicity of experimental protocols: A challenge to reductionist and non-reductionist models of the unity of neuroscience. *Synthese*, 167, 511–539.
- Sullivan, J. (2014). Stabilizing mental disorders: Prospects and problems. In H. Kincaid, & J. Sullivan (Eds.). *Classifying psychology: Mental kinds and natural kinds*. Cambridge, MA: MIT Press.
- Sullivan, J. (2016a). Construct stabilization and the unity of the mind-brain sciences. *Philosophy of Science*, 83(5), 662–673.
- Sullivan, J. (2016b). Neuroscientific kinds through the lens of scientific practice. In C. Kendig (Ed.). *Natural kinds and classification in scientific practice* (pp. 47–56). New York: Routledge.
- Sullivan, J. (2017a). Coordinated pluralism as a means to facilitate integrative taxonomies of cognition. *Philosophical Explorations*, 20(2), 129–145.
- Sullivan, J. (2017b). Long-term potentiation: One kind or many? In M. Adams, Z. Biener, U. Feest, & J. Sullivan (Eds.). *Eppur Si Muove: Doing history and philosophy of science with Peter Machamer*. Dordrecht: Springer.
- Sullivan, J. (2018). Optogenetics, pluralism and Progress. *Philosophy of Science*, 85(5), 1090–1101.
- Sullivan, J. (2020). Are there model behaviours for model organism research? Commentary on Nicole Nelson’s Model Behavior. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 82.
- Sullivan, J. (2010). Reconsidering “spatial memory” and the Morris water maze. *Synthese*, 177(2), 261–283.
- Sullivan, J. (2019). Achieving cumulative progress in understanding crime: some insights from the philosophy of science. *Psychology, Crime and Law*, 25(6), 561–576.
- Tabb, K. (2019). Philosophy of psychiatry after diagnostic kinds. *Synthese*, 196, 2177–2195.
- Tabb, K., & Schaffner, K. (2017). Causal pathways, random walks, and tortuous paths: Moving from the descriptive to the etiological in psychiatry. In K. S. Kendler, & J. Parnas (Vol. Eds.), *Philosophical issues in psychiatry. IV. Philosophical issues in psychiatry* (pp. 342–360). Oxford: Oxford University Press.
- Tsou, J. (2013). Depression and suicide are natural kinds. *International Journal of Law and Psychiatry*, 36, 461–470.
- Weiskopf, D. (2020). Anthropocentric Concepts. *Nous*, 54(2), 451–468.
- Whewell, W. (1840 [2014]). *Philosophy of the inductive sciences*. Cambridge: Cambridge University Press.
- Williams, M. (2001). *Problems of knowledge: A critical introduction to epistemology*. Oxford: Oxford University Press.
- Wimsatt, W. (2007). *Re-engineering philosophy for limited beings: Piecewise approximations to reality*. Cambridge, MA: Harvard University Press.
- Zachar, P. (2000). Psychiatric disorders are not natural kinds. *Philosophy, Psychiatry, and Psychology*, 7(3), 167–182.
- Zachar, P. (2002). The practical kinds model as a pragmatist theory of classification. *Philosophy, Psychiatry, and Psychology*, 9(3), 219–227.
- Zachar, P. (2014). *A metaphysics of psychopathology*. Cambridge, MA: MIT Press.
- Zachar, P. (2015). Psychiatric disorders: Natural kinds made by the world or practical kinds made by us? *World Psychiatry*, 14(3), 288–290.