Biological Essentialism, Projectable Human Kinds, and Psychiatric Classification

Jonathan Y. Tsou

Department of Philosophy and Religious Studies, Iowa State University

Abstract

A minimal essentialism ('intrinsic biological essentialism') about natural kinds is required to explain the projectability of human science terms. Human classifications that yield robust and ampliative projectable inferences refer to biological kinds. I articulate this argument with reference to an intrinsic essentialist account of HPC kinds. This account implies that human sciences (e.g., medicine, psychiatry) that aim to formulate predictive kind categories should classify biological kinds. Issues concerning psychiatric classification and pluralism are examined.

1. Introduction

Recently, some philosophers have resuscitated biological essentialism in the context of arguments about species. These views imply that biological taxa, including species, have essences. Ereshefsky (2010) rejects various incarnations of the 'new biological essentialism,' including the homeostatic property cluster (HPC) theory (Boyd 1999a), intrinsic essentialism (Devitt 2008, 2010), and relational essentialism (Millikan 1999; Griffiths 1999). Ereshefsky argues that there are many costs of these views (e.g., inconsistency with Darwinian theory, introducing gratuitous explanatory concepts) and no theoretical benefits. Ereshefsky's analysis fails to sufficiently appreciate an epistemic benefit of *intrinsic biological essentialism*: classifications that individuate intrinsic biological properties and mechanisms shared by kind members yield *robust and ampliative projectable inferences*.[1] Regardless of its merits for understanding species—intrinsic biological essentialism is indispensable for explaining how human science classifications produce stable projectable inferences.

The main argument of this paper is that projectable human kind classifications individuate HPC kinds constituted, in part, by intrinsic biological properties and mechanisms, i.e., an intrinsic biological essence. This argument applies to a narrow class of human sciences (e.g., medicine, psychiatry, psychology) that aim to formulate kind classifications (e.g., 'type 1 diabetes,' 'bipolar disorder,' 'type A personality') that yield stable predictions. In section 2, I defend an intrinsic

---

[1] Ampliative inferences are *informative* (i.e., non-trivial) inductive conclusions that contain content going beyond what is implicitly or explicitly contained in the premises (Salmon 1967). For inferences generated by kind terms, the content of ampliative inferences go beyond the *surface properties* ('nominal essence') used to identify kinds. A central goal of Boyd's HPC theory is to explain ampliative inferences: "[W]hat the [HPC] theory . . . helps to explain, is how we are able to identify causally sustained regularities that go beyond actually available data" (Boyd 1999a, 152). Ereshefsky (2010) mischaracterizes the HPC theory when he suggests that it aims to explain non-ampliative inferences, e.g., 'if Sparky is a dog, then Sparky will have a tail' (65). These non-ampliative inferences are also produced by *artificial kind* terms, e.g., 'if the Bible is a book, then it will contain words.'

biological essentialist account of HPC kinds, which combines insights of Boyd's HPC theory and Devitt's intrinsic essentialism. I argue that Boyd's 'accommodation thesis' is insufficient to explain the projectability of HPC kind terms and requires the further constraint that HPC kinds include *some* intrinsic properties. In section 3, I argue that projectable human kind classifications individuate HPC kinds underwritten by intrinsic biological properties and mechanisms. This argument indicates limitations of Hacking's and Mallon's analyses of human kinds. In section 4, I examine some implications of this view for psychiatric classification. I argue that intrinsic biological essentialism is consistent with a pluralistic approach to psychiatric classification and offer some amendments to Tabb's arguments for pluralism about psychiatric kinds.

2. HPC Kinds, Projectability, and Intrinsic Properties

My discussion follows a tradition initiated by Goodman (1954/1983) and expounded by Quine (1969, ch. 5), wherein a satisfactory account of natural kinds should explain how natural kind terms (or predicates) yield reliable and ampliative projectable inductive inferences. Essentialist accounts hold that natural kind terms (e.g., 'H_20') yield ampliative projectable inferences (e.g., water will freeze at 0° C) because they individuate a set of necessary and sufficient ('essential') intrinsic properties (e.g., the molecular structure of H_20) shared by kind members, which participate in laws of nature (Khalidi 2013). While essentialist accounts provide a straightforward explanation of the projectability of kind terms, they are not well-suited for describing the messier classes studied in the special sciences (e.g., biology, psychology) insofar as these classes are neither immutable nor participate in exceptionless laws of nature.

Boyd's HPC theory of natural kinds is motivated to explain the projectability of classifications in the special sciences. HPC kind terms yield projectable inferences because

members of a kind share a family of properties that cluster in a regular and non-accidental (law-like) manner. The key features of HPC kinds are the following (Boyd 1999a, 143–44):

(1) There is a family of properties (*F*) that are contingently clustered in nature.

(2) Their co-occurrence is the result of what may be described as 'homeostasis': either the presence of some properties tends to favor the presence of others, or there are underlying mechanisms that tend to maintain the properties in *F*, or both.

(3) There is a kind term (t) that is applied to things in which the homeostatic clustering of most of the properties in *F* occurs.

In Boyd's theory, the capacity of a kind term (t) to accurately represent (or accommodate) relevant causal structures—i.e., the dispositional properties and mechanisms in (2) that *cause properties to cluster in a regular and non-accidental way*—explains successful projectable inferences. According to this 'accommodation thesis,' successful inductive inferences and explanatory generalizations that are generated within a scientific framework offer (abductive) evidence that the posited kinds are accommodating (or representing) genuine causal regularities in nature. As Boyd (1991) puts it: "Kinds useful for induction or explanation must always 'cut the world at its joints' in this sense" (139).

In the literature on HPC kinds, there is disagreement regarding whether *any* of the properties or mechanisms underwriting HPC kinds need to be intrinsic. Philosophers who discuss species as HPC kinds emphasize the importance of relational properties and mechanisms (e.g., phylogenetic relations, interbreeding with conspecifics, exposure to similar environmental pressures) that maintain the stability of the property clusters associated with species (Boyd 1999a; Wilson, Barker, and Brigandt 2007). Some (Griffiths 1999; Millikan, 1999) argue that species can be explained *exclusively* in terms of relational mechanisms (e.g., descent from a common

ancestor). Others (Boyd 1999b, Wilson 1999) insist that *some* intrinsic (e.g., genetic, physiological) properties are necessary.

The central argument of this paper is that HPC kind classifications—including classifications of human kinds—must individuate *some* intrinsic properties or mechanisms to produce robust and ampliative projectable inferences. As suggested by essentialists, intrinsic properties shared by kind members explain the projectability of their classifications. In this regard, it is important to distinguish two distinctive roles that mechanisms play in the theory of HPC kinds, which are sometimes conflated in the literature:

(1) From a metaphysical standpoint, they fix the stability of natural classes and explain the unity ('homeostasis') of the observable properties that are used to identify kinds.

(2) From an epistemological standpoint, they ground the stability of projectable inferences made about such kinds.

While relational mechanisms (e.g., interbreeding with conspecifics, phylogenetic relationships) can account for (1), they are insufficient to address (2). Relational mechanisms are specified *too generally* to yield *specific inferences* about species members (cf. Devitt 2008, 2010). For projectable inferences about a particular species to be made, a kind term needs to individuate *some* intrinsic properties (e.g., genetic properties) and internal mechanisms (e.g., developmental mechanisms) common to kind members.

The argument that projectable HPC kind classifications must individuate intrinsic properties contrasts with Boyd's approach. In response to an argument by Millikan (1999) that species are 'historical kinds' constituted entirely by relational mechanisms (e.g., a copying process), Boyd (1999b) objects that species are HPC kinds underwritten by *both* intrinsic (e.g., genetic properties, phenotypic traits) and relational mechanisms. The importance of intrinsic

properties is that these "(imperfectly) shared properties" explain the projectability of HPC terms (Boyd 1999b, 81–2). Elsewhere, Boyd (1991, 1999a) suggests that some HPC kinds can be defined almost entirely in terms of relational and conventionally stipulated properties, and he presents 'social roles' and 'feudal economy' as examples. Boyd's neutrality on the properties individuated by HPC kind terms reflects his *a posteriori* methodological stance that the 'naturalness' of natural kinds is *whatever* reference to such kinds contributes to the accommodation of classificatory schemes to causal structures (Boyd 1999a, 158–9). If the consideration he raises in response to Millikan (i.e., projectable kinds need to individuate some intrinsic properties) is relevant for species, it should apply to HPC kinds more generally. Hence, the 'naturalness' of natural kinds should be identified—as in traditional essentialist accounts—with intrinsic natural properties.

3. Human Kinds, Stability, and Projectability

Human kind classifications yield robust and ampliative projectable inferences when they individuate *biological kinds*, i.e., HPC kinds constituted (in part) by a set of intrinsic biological properties and mechanisms. This perspective suggests limitations of arguments made by Ian Hacking and Ron Mallon. Contra-Hacking, human classifications that individuate biological kinds are sufficiently stable to yield robust projectable inferences despite their production of feedback effects. Contra-Mallon, social kinds constituted exclusively by relational mechanisms may yield transitory projectable inferences, but they will not yield the robust projectable inferences sought by human sciences (e.g., cognitive psychology, medicine) that aim to formulate *predictive* categories.

An issue raised in the Hacking-Boyd exchange at the 1989 Oberlin Colloquium (Hacking 1991a, 1991b; Boyd 1991) is whether human or social kinds can be natural kinds. While Boyd is

optimistic that the HPC theory can apply to human kinds, Hacking (1991b) is pessimistic, arguing that—in contrast with natural kind classifications—human kind classifications (e.g., 'multiple personality,' 'homosexuality') are 'made up' insofar as "the histories of human kinds are . . . different from those of natural kinds, for as [human] classifications are formed and moulded they loop back, interact with, and alter the individuals and the types of behaviour to which they apply" (152–3). This is a prototype of what Hacking (1995) subsequently formulates as the 'looping effects of human kinds': the meaning of human science classifications *changes* classified people *such that the classifications need to be constantly revised*. While the natural kinds (or 'indifferent kinds') classified in the natural sciences (e.g., 'electrons,' 'H_2O') will not change based on how they are classified, the human kinds (or 'interactive kinds') classified in the human and social sciences (e.g., 'attention deficit hyperactivity disorder,' 'autism') are aware of and will change in response to how they are classified (Hacking 1999; cf. Khalidi 2010). Because of looping effects, the targets of classification in the human sciences are "on the move" (Hacking 1999, 108). Hacking (2007) writes:

> We think of … kinds of people … as definite classes defined by definite properties. … But it is not quite like that. They are moving targets because our investigations interact with the targets themselves, and change them. And since they are changed, they are not quite the same kind of people as before. The target has moved. That is the looping effect. (293)

Since human science classifications change the people they classify, classified people *constantly change* in response to how they are classified and their classifications need to be *constantly revised* to accommodate these changes. If human kinds and their classifications are *perpetually unstable*,

human kinds cannot be natural kinds and their classifications cannot yield stable projectable inferences.

Hacking's conclusion that objects of human science classifications are inherently unstable ('moving targets') is a non-sequitur that does not follow from his (correct) claim that human science classifications change classified people (Tsou 2021). In particular, Hacking conflates and fails to distinguish two separate theses:

(1) *Classificatory feedback*: Human science classifications change the experiences and behavior of the people they classify.

(2) *Looping Effects*: Human science classifications must be constantly revised to accommodate changes they cause among classified people.

Classificatory feedback is a ubiquitous feature of human science classifications. If classified people are aware of and will change in response to how they are classified, then (1) will be present. However, the presence of (1) does not entail (2), which is the key premise supporting Hacking's conclusion that the targets of human science classifications are inherently unstable. If a human science classification refers to a biological kind, then (1) will not render that object of classification unstable. This applies to medical classifications (e.g., 'HIV,' 'breast cancer') that cause classificatory feedback. While being diagnosed as HIV positive will inevitably change the experiences and behavior of classified individuals, such changes will not require revising the meaning of HIV nor the symptoms used to diagnose it. The HIV classification is stable because it refers to a biological kind (viz., an immunodeficiency virus). This also applies to psychiatric classifications (e.g., 'depression,' 'schizophrenia') that individuate biological kinds. While being diagnosed with 'depression' will undoubtedly change classified individuals, these changes will not require revising the meaning of 'major depression' or the criteria used to define it (e.g., persistent

feelings of sadness). Hence, biological kinds are stable targets of classification despite the presence of classificatory feedback. While there are instances of looping effects in the human sciences (e.g., 'hysteria,' 'multiple personality'), these cases are far less prevalent than Hacking suggests and involve classifications that primarily track social kinds, rather than biological ones.

The argument that projectable human classifications individuate biological kinds implies that classifications of social kinds cannot yield *robust* projectable inferences. Drawing on Boyd's enthusiasm for extending HPC kinds to social kinds, Mallon (2016) argues that socially constructed human classifications (e.g., 'race,' 'gender') yield projectable inferences when they refer to HPC kinds constituted by social properties. Mallon defends a realist account of social construction, which explains the (social) reality of human kinds in terms of 'social roles.' Social roles are structured by a variety of *representations* (e.g., attitudes, theories, narratives, models, norms, texts). A social role exists when: (1) representations pick out a category of persons and a set of beliefs and evaluations associated with the term, and (2) most of the beliefs and evaluations associated with the role are common knowledge (Mallon 2016, 6). Condition (2) is achieved through the public broadcasting of information (e.g., by scientific institutions or the media) about the social role. When a social role (e.g., 'woman') is *entrenched* by various social (relational) mechanisms (e.g., looping effects that stabilize a category), they are transformed into *human kinds*: social categories that support inductive inferences, predictions, and explanation. Mallon's explanation of the projectability of socially constructed human categories appeals to the idea that *entrenched social roles* (or 'human kinds') are HPC kinds (Mallon 2016, 89). Mallon finds Boyd's HPC theory apt for characterizing social roles given its liberalization of the properties that constitute kinds to include relational properties, in addition to intrinsic ones (cf. Boyd 1991, 140–3). This "opens the door to properties that result from human practices, norms, conventions, and

so forth" (Mallon 2016, 92). If entrenched social roles are HPC kinds, then this could explain the projectability of socially constructed human kinds:

> [C]ausally significant social roles . . . could figure as the homeostatic mechanism at the center of important property cluster-kinds that structure our social world. . . [S]ocial roles that . . . produce and sustain property-cluster kinds may support induction, prediction, and explanation. (Mallon 2016, 92)

Mallon's account of how social roles *produce* HPC kinds is somewhat murky, but the basic idea is that entrenched social roles acquire *causal significance* insofar as there are real and systematic causal consequences of being member of these constructed categories, which ground projectable inferences associated with their classifications.

While the social kinds invoked in Mallon's social role account may yield transitory projectable inferences, human classifications that individuate biological kinds yield stable and robust projectable inferences (Tsou 2020). Consider Mallon's paradigms of social kinds: race and gender. Mallon is interested in how these categories yield projectable inferences related to oppression. In this light, these categories currently yield some ampliative inferences (e.g., African American men are more likely to be arrested by the police in the United States, women will earn less money than men). These projectable inferences are *transitory* because the social mechanisms that ground them (e.g., racism, sexism) are subject to change. By contrast, classifications of biological kinds (e.g., 'PKU') yield robust projectable inferences (e.g., placing PKU infants on a phenylalanine-free diet during a critical period allows brain myelination to occur normally) over longer time-scales and across different social contexts. These projectable inferences are robust because they are *causal consequences* of stable intrinsic biological properties (e.g., PKU is caused by a lack of a liver enzyme, which interferes with brain myelination when phenylalanine enters

the bloodstream). Hence, projectable inferences yielded by classifications of biological kinds will be as stable as the biological processes that constitute such kinds. For diseases, this assumes that there are species-specific ('normal') functions that contribute to health and interferences to these functions result in systematic causal effects (Boorse 1977). While biological properties and mechanisms are subject to change due to evolutionary processes, they are sufficiently stable to ground robust projectable inferences. This perspective is amenable to Woodward's view that explanatory generalizations in the special sciences should be *invariant* under various interventions and background conditions (Woodward 2000). While generalizations about social kinds will vary across different social contexts, generalizations about biological kinds have a much wider domain of invariance.[2]

4. Psychiatric Kinds, Pluralism, and Levels of Classification

In the following section, I examine issues of psychiatric classification in light of the analysis of projectability defended in sections 2 and 3. While projectable human science classifications individuate biological kinds, there is no privileged or uniquely correct way of classifying biological kinds. Human science classifications, like all scientific classifications, involve elements of convention since they individuate classes that are relevant to the values and interests of classifiers. This perspective draws on Dupré's stance on natural kinds ('promiscuous realism') that accommodates aspects of realism and pluralism: "The realism derives from the fact that there are many sameness relations that serve to distinguish classes … in ways that are relevant to various

---

[2] My analysis does not imply that research on social kinds is illegitimate. Rather, it targets a subset of human sciences—of which medicine and psychiatry are paradigms—that aim to formulate (law-like) human kind categories *assumed to be stable natural kinds*. Standards of robustness for projectable inferences will vary between disciplines and with the aims of researchers.

concerns; the promiscuity derives from the fact that none of these relations is privileged" (Dupré 1981, 82). A pluralistic approach to kinds and classification implies that there are multiple legitimate levels of classification, so long as classifications pick out *natural similarities* that unify members of a class. My analysis of projectability implies that *predictively useful* human classifications individuate kinds constituted by intrinsic biological properties and mechanisms.

In contemporary debates about psychiatric classification, there is disagreement regarding the appropriate *level* to classify psychiatric constructs. The most influential system, the *Diagnostic and Statistical Manual of Mental Disorder* (DSM), classifies constructs at the level of 'mental disorders,' which are assumed to be discrete disease syndromes identifiable at the individual and population level. In the DSM, mental disorders are defined operationally by a polythetic set of necessary and sufficient diagnostic criteria. The DSM has been criticized on numerous grounds, including its inability to formulate valid diagnostic categories, the high co-occurrence ('comorbidity') of categories (e.g., depression and anxiety) that are assumed to be causally distinct, and its failure to adequately incorporate empirical findings (Tsou 2021). Herein, I focus on the methodological criticism that the DSM formulates its classifications at an inappropriate level of analysis.

Disenchantment with the failure of the DSM to provide *valid* diagnostic categories (i.e., classifications that accurately represent real phenomena) stimulated the formulation of alternative classifications systems that define psychiatric constructs at different levels of analysis. The Research Domain Criteria (RDoC) classifies psychiatric constructs (e.g., 'auditory perception,' 'potential threat/ anxiety,' 'reward prediction error') at lower levels of analysis than the DSM and explicitly demands that there are biological correlates (e.g., genetic or neurobiological biomarkers) for these constructs (see NIMH 2018). In contrast with the DSM's categorical approach to

classification, the Hierarchical Taxonomy of Psychopathology (HiTOP) adopts a dimensional approach that organizes psychiatric constructs in a hierarchical manner with *spectrum constructs* (e.g., 'internalizing,' 'externalizing') at the top of the hierarchy, which subsume more specific *domain constructs* (e.g., 'fear,' 'distress'), which subsume *disorders* (e.g., 'depression'), which subsume *behaviors* and *symptoms* at the bottom of the hierarchy (Kotov et al. 2017, 462).

Drawing on methodological criticisms of the DSM by advocates of these alternative systems, Kathryn Tabb (2015) argues against the DSM's assumption of 'diagnostic discrimination,' which maintains that "the operationalized criteria for diagnosing clinical types will . . . successfully pick out populations about which relevant biomedical facts can be discovered" (1048). This argument emphasizes the polythetic nature of DSM categories, which generate an immense array of incongruent symptom-profiles. Tabb argues that the *heterogeneous* groups of patients produced by DSM definitions are unsuitable for discovering underlying mechanisms. She praises alternative systems for offering novel ways of formulating psychiatric constructs and compiling research populations that do not rely on DSM definitions. Tabb (2019) criticizes philosophers of psychiatry for implicitly accepting the DSM assumption that psychiatric constructs formulated at the level of mental disorders ('diagnostic kinds') are the most important psychiatric kinds to investigate. Against this view, she defends a pluralistic view that encourages investigating a multiplicity of psychiatric kinds (e.g., risk factors, personality dimensions, phenomenological states, social conditions) other than diagnostic kinds. In Tabb's ideal of integrative pluralism (cf. Mitchell 2003), efforts to integrate a metaphysically diverse array of psychiatric kinds provides a more promising path for identifying valid and pragmatically useful psychiatric categories.

Tabb's argument that DSM definitions have impeded the formulation of valid psychiatric categories in research contexts is compelling; however, she overstates the impossibility of

diagnostic discrimination. For example, since the mid-1990s, the DSM has defined schizophrenia by the following symptoms (APA 2013, 99-100):

(1) delusions

(2) hallucinations

(3) disorganized speech

(4) grossly disorganized or catatonic behavior

(5) flat emotional expression, poverty of speech, or lack of motivation

Despite content validity problems with this definition (e.g., the exclusion of cognitive impairments that are common in schizophrenia), it has been *sufficiently valid* to facilitate the discovery of novel neurobiological mechanisms (e.g., glutamate systems, impaired neuroplasticity) and genetic biomarkers (e.g., genetic anomalies on chromosomes 6 and 1) in genome-wide association studies (see Tsou 2021). This counters Tabb's pessimism that diagnostic kinds *cannot* produce research samples suitable for discovering biomedical facts.

The analysis of this paper suggests that both the validity and projectability of psychiatric classifications requires the accurate individuation of stable biological kinds. Elsewhere, I have argued that multiple lines of independent research (e.g., pharmacology, brain imaging studies, animal models) support the inference that the DSM's schizophrenia classification tracks a stable and complex neurobiological kind (Tsou 2012, 2021). In this regard, the specific symptoms used to define schizophrenia correlate with distinctive neurobiological biomarkers (or criterion-related validators). For example, the positive (or 'psychotic') symptoms of schizophrenia (symptoms 1 and 2) correlate with excessive dopamine activity in the mesolimbic pathway, while negative symptoms (symptom 5) correlate with deficient dopamine activity in the mesocortical pathway and deficient glutamate activity in the prefrontal cortex. Moreover, negative and positive

14

symptoms tend to co-occur because deficient activity of dopamine in the prefrontal cortex (the terminal region of the mesocortical pathway) correlates with excessive dopamine activity in mesolimbic pathway.[3] Robust and ampliative projectable inferences yielded by the schizophrenia classification (e.g., schizophrenia can be treated by dopamine antagonist drugs or glutamate agonist drugs) are explained by the fact that it reliably tracks these intrinsic biological properties.

While Tabb's integrative pluralism is unobjectionable, her analysis should be qualified by the demand that psychiatric kinds formulated at various levels of analysis should individuate (or causally relate to) biological kinds. This prescription is consonant with the RDoC's bottom-up strategy of formulating psychiatric constructs that correlate with biomarkers at different (e.g., molecular, cellular, neurobiological) levels. For example, the RDoC construct of 'potential threat/ anxiety' correlates—at the level of neurobiological circuitry—with activity in the bed nucleus of the stria terminalis (NIMH 2018). Compared to the 'diagnostic kind model' adopted by the DSM (Tabb 2019), the strategy of correlating hypothesized psychiatric constructs with biomarkers provides a more direct and transparent approach for ensuring that psychiatric constructs possess (criterion-related) validity. The HiTOP strategy of classifying psychiatric constructs at different hierarchical levels, wherein higher-level constructs subsume lower-order constructs, suggests that some DSM categories (e.g., 'major depression,' 'generalized anxiety,' 'PTSD') can be grouped together at a higher level (e.g., 'distress disorders'). Hyman (2007) points out that the architects of DSM-III somewhat arbitrarily decided to 'split' (rather than 'lump') disorders, which resulted in a proliferation of diagnostic categories in subsequent editions. My analysis suggests that disorders

---

[3] My analysis is neutral on the *causal priority* of different biological mechanisms and maintains that various mechanisms (e.g., dopamine, glutamate, serotonin, GABA, impaired neuroplasticity) *interact* to produce the signs of schizophrenia. This assumes that the glutamate hypothesis is consistent with the dopamine hypothesis and motivates an expansion of the latter, rather than a rejection of it (Tsou, 2021).

should be classified at a level that individuates a *causally distinctive biological kind* (cf. Craver 2009), which implies that fine-grained DSM categories (e.g., 'obsessive-compulsive disorder' 'agoraphobia,' 'social phobia') should be classified ('lumped') at a higher level of analysis (e.g., 'fear disorders') if they are constituted by similar biological mechanisms. Elsewhere, I have argued that these fine-grained psychiatric categories (and 'culture-bound syndromes') are differentiated by the social mechanisms emphasized by Hacking and Mallon (e.g., broadcasting of social roles, imitation of stereotypes), which *stabilize a culturally-specific expression* of a biological kind (Tsou 2021).

5. Conclusion

While intrinsic biological essentialism may be dispensable for formulating species concepts, it is indispensable for understanding the predictive inductive projects pursued in human sciences such as medicine and psychiatry. Human science classifications that track a stable set of intrinsic biological properties ('biological kinds') yield robust and ampliative projectable inferences. Against Hacking's analysis, biological kinds are stable objects of human classification despite the feedback effects generated by their classifications. As a qualification about Mallon's analysis, classifications of social kinds can yield transitory projectable inferences, but they cannot yield the robust inferences sought in human sciences like medicine or psychiatry. As indicated in my analysis of psychiatric classification, intrinsic biological essentialism is consistent with pluralistic approach to kinds and classification. To facilitate the formulation of valid and predictive psychiatric constructs, Tabb's integrative pluralism should be constrained by the requirement that psychiatric constructs formulated at various levels correspond (or relate) to causally distinctive biological kinds.

References

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

Boorse, Christopher. 1977. "Health as a Theoretical Concept." *Philosophy of Science* 44(4): 542–73.

Boyd, Richard. 1991. "Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds." *Philosophical Studies* 61(1–2): 127–48.

Boyd, Richard. 1999a. "Homeostasis, Species, and Higher Taxa." In Wilson 1999b, 141–85.

Boyd, Richard N. 1999b. "Kinds, Complexity, and Multiple Realization: Comments on Millikan's 'Historical Kinds and the Special Sciences.'" *Philosophical Studies* 95(1–2): 67–98.

Craver, Carl F. 2009. "Mechanisms and Natural Kinds." *Philosophical Psychology* 22(5): 575–94.

Devitt, Michael. 2008. "Resurrecting Biological Essentialism." *Philosophy of Science* 75(3): 344–82.

Devitt, Michael. 2010. "Species Have (Partly) Intrinsic Essences." *Philosophy of Science* 77(5): 648–61.

Dupré, John. 1981. "Natural Kinds and Biological Taxa." *Philosophical Review* 90(1): 66–90.

Ereshefsky, Marc. 2010. "What's Wrong with the New Biological Essentialism?" *Philosophy of Science* 77(5): 674–85.

Goodman, Nelson. 1954/1983. *Fact, Fiction, and Forecast*, 4th ed. Cambridge, MA: Harvard University Press.

Griffiths, Paul E. 1999. "Squaring the Circle: Natural Kinds with Historical Essences." In Wilson 1999b, 209–28.

Hacking, Ian. 1991a. "A Tradition of Natural Kinds." *Philosophical Studies* 61(1–2): 109–26.

Hacking, Ian. 1991b. "On Boyd." *Philosophical Studies* 61(1–2): 149–54.

Hacking, Ian. 1995. "The Looping Effects of Human Kinds." In *Causal Cognition: A Multidisciplinary Debate*, eds. Dan Sperber, David Premack, and Anne J. Premack, 351–83. Oxford: Clarendon Press.

Hacking, Ian. 1999. *The Social Construction of What?* Cambridge, MA: Harvard University Press.

Hacking, Ian. 2007. "Kinds of People: Moving Targets." *Proceedings of the British Academy* 151: 285–318.

Hyman, Steven E. 2007. "Can Neuroscience be Integrated into the DSM-V?" *Nature Reviews Neuroscience* 8(9): 725–32.

Khalidi, Muhammad Ali. 2010. "Interactive Kinds." *British Journal for the Philosophy of Science* 61(1): 335–60.

Khalidi, Muhammad Ali. 2013. *Natural Categories and Human Kinds: Classification in the Natural and Social Sciences*. Cambridge: Cambridge University Press.

Kotov, Roman, Robert F. Krueger, David Watson, et al. 2017. "The Hierarchical Taxonomy of Psychopathology (HiTOP): A Dimensional Alternative to Traditional Nosologies." *Journal of Abnormal Psychology* 126(4): 454–77.

Mallon, Ron. 2016. *The Construction of Human Kinds*. Oxford: Oxford University Press.

Millikan, Ruth G. 1999. "Historical Kinds and the 'Special Sciences.'" *Philosophical Studies* 95(1–2): 45–65.

Mitchell, Sandra D. 2003. *Biological Complexity and Integrative Pluralism*. Cambridge: Cambridge University Press.

NIMH. 2018. RDoC Matrix Archives. National Institute of Mental Health. ([link](#))

Quine, W. V. 1969. *Ontological Relativity and Other Essays*. New York: Columbia University Press.

Salmon, Wesley C. 1967. *The Foundations of Scientific Inference*. Pittsburgh: University of Pittsburgh Press.

Tabb, Kathryn. 2015. "Psychiatric Progress and the Assumption of Diagnostic Discrimination." *Philosophy of Science* 82(5): 1047–58.

Tabb, Kathryn. 2019. "Philosophy of Psychiatry after Diagnostic Kinds." *Synthese* 196(6): 2177–99.

Tsou, Jonathan Y. 2012. "Intervention, Causal Reasoning, and the Neurobiology of Mental Disorders: Pharmacological Drugs as Experimental Instruments." *Studies in History and Philosophy of Biological and Biomedical Sciences* 43(2): 542–55.

Tsou, Jonathan Y. 2020. "Social Construction, HPC Kinds, and the Projectability of Human Categories." *Philosophy of the Social Sciences* 50(2): 115–37.

Tsou, Jonathan Y. 2021. *Philosophy of Psychiatry*. Cambridge: Cambridge University Press.

Wilson, Robert A. 1999a. "Realism, Essence, and Kind: Resuscitating Species Essentialism?" In Wilson 1999b, 187–207.

Wilson, Robert A., ed. 1999b. *Species: New Interdisciplinary Essays*. Cambridge, MA: MIT Press.

Woodward, James. 2000. "Explanation and Invariance in the Special Sciences." *British Journal for the Philosophy of Science* 51(2): 197–254.