Forthcoming in *The Monist* (2025)

Hacking on Looping Effects and Kinds of People

Jonathan Y. Tsou
University of Texas at Dallas

ABSTRACT

This paper critically examines Ian Hacking's account of looping effects and human kinds, focusing on three related arguments defended by Hacking: (1) the looping effects of human science classifications render their objects of classification inherently unstable, (2) looping effects preclude the possibility of generating stable projectable inferences (i.e., reliable predictions) based on human kind terms, and (3) looping effects can demarcate human science classifications from natural science classifications. Contra-Hacking, I argue that: (1) some objects of human science classifications (viz., biological kinds) remain stable despite the feedback generated by their classifications, (2), human science classifications that individuate biological kinds yield stable projectable inferences, and (3) looping effects are a problematic criterion for distinguishing human science classifications from natural science classifications.

1. Introduction

Two enduring ideas from Ian Hacking's corpus are making up people and the looping effects of human kinds. Hacking's making up people project examines the way that human science classifications (e.g., 'multiple personality disorder,' 'autism,' 'homosexuality') create new *kinds of people* (i.e., new ways for people to *understand and conceive of themselves*) that did not exist before they were named. Looping effects are the main social mechanism that explains the interactive causal relationship between human science classifications and the classes ('kinds') of people that they classify. Looping effects are a causal feedback loop, whereby the *meaning of a human science classification* causes changes to classified people, which subsequently requires revisions to the classification (Hacking 1995b). The stereotypical *meaning* of human science classifications (e.g., 'major depressive disorder') *changes the experiences and behavior of the people classified* (e.g., people diagnosed with depression act in accordance with prevailing stereotypes). The ways in which classes (or subclasses) of people change in response to being classified (e.g., depression is expressed in terms of guilt) can demand corresponding revisions to the classification. In contrast to natural science classifications (e.g., 'quarks,' 'electrons,' 'sulfur'), human science classifications loop back on themselves via feedback from the classes ('kinds') of people being classified.

This paper critically examines Hacking's account of looping effects and human kinds. My analysis focuses on three related arguments that Hacking has defended: (1) the looping effects of human science classifications render their objects of classification inherently unstable or 'moving targets,' (2) looping effects preclude the possibility of generating stable projectable inferences (i.e., reliable predictions) based on human kind terms, and (3) looping effects can stand as a demarcation criterion for distinguishing the human and natural sciences. Contra-

Hacking, I argue that: (1) some objects of classification in the human sciences (e.g., natural kinds) can remain stable despite the feedback generated by their classifications, (2), some human science classifications can yield stable projectable inferences in spite of classificatory feedback, and (3) looping effects are a problematic criterion for distinguishing the human sciences from the natural sciences.

The paper proceeds as follows. In section 2, I explicate Hacking's account of looping effects, focusing on his argument that looping effects render their objects of classification inherently unstable. This general argument motivates Hacking's more specific arguments regarding the projectability of human science classifications and demarcating the human sciences from the natural sciences. In section 3, I argue that Hacking's argument that looping effects render human kinds inherently unstable is a hasty generalization that conflates classificatory feedback (i.e., human science classifications constantly change the experiences and behavior of classified people) with looping effects (i.e., human science classifications must constantly be revised to account for ways in which classified people are changed by classifications). In section 4, I argue that human science classifications of biological kinds (e.g., 'introversion,' 'breast cancer,' 'schizophrenia') yield reliable projectable inferences. I support this argument by articulating an account of biological kinds that combines insights from Richard Boyd's homeostatic property cluster theory of natural kinds and Michael Devitt's arguments for intrinsic essentialism. In section 5, I offer reasons for rejecting looping effects as a criterion for distinguishing the human sciences from natural sciences. At the outset, it should be stated that my aim is not to deny the existence of looping effects, but to challenge the ubiquity and frequency of looping effects in the human sciences as presented by Hacking.

2. Looping Effects and the Instability of Human Kinds

Hacking argues that the kinds and classifications formulated in the human and social sciences ('human kinds,' 'kinds of people,' or 'interactive kinds') are fundamentally different from the natural kind classifications formulated in natural sciences, such as physics and chemistry. My presentation focuses on Hacking's later analyses (Hacking 1999, 2002b, 2007), wherein Hacking *identifies* the reactivity of human subjects (i.e., people's awareness and responsiveness to being classified) with looping effects, and he argues that the targets of human classifications are *inherently unstable*. In his seminal paper on looping effects, Hacking (1995b) provides a more nuanced and qualified approach to human kinds. Therein, he presents 'human kinds' as a *particularly socially-constituted* type of human science classifications (1995b, 353), which he restricts to human science classifications (e.g., 'child abuse,' 'multiple personality') that are peculiar to *a specific social setting* (i.e., a particular historical time and geographical location). By contrast, *natural* human science classifications (e.g., 'mass,' 'longevity,' 'digestive organs,' 'the structure of the genome') are understood as referring to natural kinds that appear ('universally') across different social settings (cf. Hacking 1991a, 123). In later works, Hacking (1999, 2002b, 2007) no longer mentions this distinction (cf. Hacking 1999, 104-23). One motivation of this paper (section 4) is to articulate a defensible version of Hacking's early distinction between *social* and *natural* human science classifications.[1]

[1] The main ideal that Hacking (1995b) aims to undermine—in his early and late works—is the ideal of *projectable* human kind classifications: "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" (353). My arguments in sections 3 and 4 are intended to show that the scope of projectable human kinds is much larger than Hacking's list of human natural kinds (Hacking 1995b, 253). In defending the possibility of projectable human kind classifications, I follow Hacking's focus on human sciences (e.g., medicine, psychology, psychiatry) *that aim to provide predictive classifications* (see Hacking 1995b, 2007, 93), My analysis does not apply to human sciences (e.g., anthropology, sociology, history) that do not primary aim to formulate *predictive* categories.

A distinctive feature of objects of classification in the human sciences is that the people being classified are *aware of and will change in response to how they are classified*.[2] While objects of classification in the natural sciences (e.g., electrons, water) are *indifferent* to how they are classified, objects of classification in the human sciences (e.g., people with schizophrenia, children with autism) *interact* with their classifications (Hacking 1999, ch. 4). This feature of classified people gives rise to the looping effects of human kinds: the meaning of human science classifications (e.g., 'multiple personality') constantly change the *experiences and behaviors of classified people* (e.g., individuals act in accordance with prevailing stereotypes) such that the classifications must be constantly revised to accommodate such changes. Hacking emphasizes that his account of looping makes no specific predictions about how classes of people will react to being classified (e.g., conformity versus non-conformity), but it implies that the 'kinds of people' individuated by human science classifications will *continuously evolve and require classificatory revisions* to capture such changes: "*People classified in a certain way tend to conform to or grow into the ways that they are described*; but they also evolve in their own ways, so that the classifications and descriptions have to be constantly revised" (Hacking 1995a, 21).[3]

Hacking's argument that human science classifications *change* classified people motivates his further claim that the objects of classification in the natural sciences are stable, while the objects of classification in the human sciences—because of looping effects—are *inherently unstable*. Hacking (2007) writes:

---

[2] Recently, some philosophers of science have discussed this feature as the 'reactivity' of human subjects (e.g., see Runhardt 2021, 2022; Fagerman 2022; Marchionni, Godman, and Zahle 2024). In section 5 of this paper, I argue that reactivity is not limited to objects of study in the human sciences and it can apply to some objects of study in the natural sciences (e.g., domesticated kinds).

[3] For critical discussion of Hacking's suggestion (see Hacking 1984, 48-60; 1986, 108-110; 1995b, 68-9; 1999, 103-105) that 'kinds of people' can be understood in terms of Elizabeth Anscombe's theory that intentional actions are 'actions under a description' (Anscombe, 1957, 1979), see Cooper (2004, 80-4) and Allen (2021, S2930-34).

> 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 [and classifications] 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)

In *The Social Construction of What?*, Hacking (1999) suggests that the presence or absence of looping effects offers a criterion for distinguishing between classifications in the human sciences (i.e., 'interactive kinds' that interact with their objects of classification) and classifications in the natural sciences (i.e., 'indifferent kinds' that do not interact with their objects of classification because their objects of classification are *indifferent* to how they are classified):

> [A] cardinal difference between the traditional natural and social sciences is that the classifications employed in the natural sciences are indifferent kinds, while those employed in the social sciences are mostly interactive kinds. The targets of the natural sciences are stationary. Because of looping effects, the targets of the social sciences are on the move. (108)

In these passage, Hacking draws the radical conclusion that—because human science classifications *change* the people they classify—the objects of human science classifications ('kinds of people') are *perpetually unstable entities*. Since the meaning of human science classifications constantly change the people they classify, *classified people will constantly change* in response to how they are classified and *classifications will need to be constantly revised* to accommodate such changes. This is the precise sense in which Hacking argues that looping effects render their objects of classification 'moving targets.' One surprising implication of this analysis is that human science classifications *cannot* yield any stable projectable

inferences (i.e., predictions about members of a kind) given the inherent instability of the objects that they classify (Hacking 1995b).

3. Classificatory Feedback Does Not Imply Instability

Hacking commits a hasty generalization in concluding that the objects of classification in the human sciences are inherently unstable (Tsou 2007, 2013, 2016, 2021). In particular, Hacking's conclusion is a non-sequitur that does not follow from his (correct) claim that human science classifications change the people they classify. Hacking's argument can be reconstructed as follows:

(1) People are aware of and will change in response to how they are classified (*the reactivity of human subjects).*
(2) Human science classifications change the experiences and behavior of the people they classify (*classificatory feedback*).
(3) Human science classifications must be revised in order to accommodate changes that the classifications cause among classified people (*looping effects*).
(4) Thus, objects of classification in the human sciences are inherently unstable (*human kinds are moving targets*).

While (1) and (2) are undoubtedly true, Hacking fails to demonstrate that these premises entail (3), which is the key premise that establishes (4).[4] Tsou (2007) argues that Hacking's analysis conflates (2) and (3).[5] Classificatory feedback is a *ubiquitous* feature of most human science

[4] My argument assumes that (3) is universally (rather than existentially) quantified. My own view is that (3) should be existentially quantified. One could argue that Hacking also assumes that (3) is existentially quantified (cf. Hacking 1995b). My justification for presenting Hacking as endorsing a universal interpretation of (3) is that it is needed to support his generalizations that: (1) the targets of the natural sciences are stationary while those in the social sciences are constantly moving (Hacking 1999, 108), and (2) kinds of people are moving targets (Hacking 2007, 293).

[5] What Tsou (2007) calls 'weak implications of looping' is referred to herein as 'classificatory feedback'; what he calls 'strong implications of looping' is referred to as 'looping effects.' In early work, Hacking (1986, 111-112) distinguishes between a 'vector from above' (experts labeling people) and 'vector from below' (autonomous behavior of labeled people), which roughly corresponds to the distinction between classificatory feedback and looping effects. Hacking does not refer to this distinction in subsequent works (cf. Hacking 1998, 80-87; Hacking 2007, 305-310).

classifications. So long as the people being classified are aware of and will change in response to how they are classified, classificatory feedback will be present.[6] However, the mere presence of classificatory feedback does not necessarily imply looping effects. While diagnosing an individual as 'depressed' will invariably change the experiences and behavior of the individual, the ways in which classificatory feedback change groups (or sub-groups) of classified people will not necessarily require corresponding changes to the *meaning* of 'clinical depression,' revisions in the criteria used to identify depression (e.g., persistent feelings of sadness), or our knowledge about these classes of people.

For looping effects to occur, very specific conditions must be met. Classificatory feedback must change the experiences and behavior of individuals—*in a uniform manner for a significant number of individuals who fall under that classification*—such that the *meaning* of the classification (e.g., the criteria that constitute membership for a classification, stereotypes of the classification) must be revised. Depending on how many individuals are *uniformly changed*, classificatory feedback might lead to the creation of subtypes of classification. In cases where a human science classification refers to a biological natural kind—a circumstance that I discuss in section 4—classificatory feedback may not require any revisions to the classification. While there are examples of looping effects in the human sciences (e.g., 'hysteria,' 'multiple personality'), their prevalence is far less frequent than Hacking suggests. The identification of looping effects must be made on a case-to-case basis, rather than in an *a priori* manner.

Against Hacking's contention that the objects of human science classifications are inherently unstable 'moving targets' *because of looping effects*, the targets of some human

[6] Tekin (2011, 2014) offers a rich elaboration and expansion of classificatory feedback by focusing on concepts of the self and narratives, which are underdeveloped in Hacking's account.

science classifications are stable objects of classification (i.e., biological kinds) *despite the presence of classificatory feedback*. If a human science classification (e.g., 'introversion,' 'schizophrenia,' 'HIV') individuates a (relatively) stable biological kind, then the presence of classificatory feedback generated by that classification will not necessarily render its object of classification (viz., a biological kind) unstable nor render the classification incapable of producing stable projectable inferences.

4. Human Classifications of Biological Kinds are Projectable

One of the primary motivations of Hacking's original account of looping effects was to argue against the possibility of formulating human science classifications modeled on the natural sciences and especially the tendency to biologize human kinds (Hacking 1995b, 353-5, 362-4, 372-3). As Hacking (1995b) puts it: "biologizing human kinds does not thereby make them immune to looping effects" (372). On Hacking's view, looping effects preclude the possibility of formulating human science classifications that individuate *stable* natural kinds, which provide "systematic, general, and accurate knowledge . . . that could be used to formulate generalizations . . . precise enough to predict what individuals will do, or how they will respond to attempts to help them or to modify their behaviour" (Hacking 1995b, 352). In the terminology of *projectability* (Goodman 1954/1983; Quine 1969, ch. 5; Boyd 1991; Millikan 1999), looping effects preclude the possibility of formulating *projectable* human science classifications, i.e., classifications that yield reliable and informative (i.e., ampliative) inferences about members of a class on the basis of induction.[7]

[7] Ampliative inferences are informative (i.e., nontrivial) 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

Contra-Hacking, human science classifications (e.g., 'HIV,' 'breast cancer,' 'PKU,' 'introversion,' 'Type A,' 'schizophrenia' 'bipolar disorder') that refer to *biological kinds* yield reliable projectable inferences. Generally, biological kinds are relatively stable classes (e.g., of physical signs, behaviors, psychological dispositions) that are *definable* by biological properties, mechanisms, or biomarkers. My analysis does not imply that 'biological kinds' are value-free or free from social elements. I assume that: (1) the distinction between natural and artificial kinds marks a distinction between *discovered classes* (e.g., 'electrons,' 'oxygen,' 'fish') and *invented classes* (e.g., 'chairs,' 'skyscrapers,' 'games') respectively, and (2) this distinction is a *distinction of degree*. While *all scientific classifications* (including human science classifications) involve elements of (social) invention insofar as they individuate classes that are *relevant* to the values and interests of classifiers, some classifications are *more natural* insofar as they individuate classes in terms of (discovered) *natural properties*. This interest-relative view on natural kinds draws on Dupré's "promiscuous realism," which accommodates realist and pluralist commitments: "The realism derives from the fact that there are many sameness relations that serve to distinguish classes . . . in ways that are relevant to various concerns; the promiscuity derives from the fact that none of these relations is privileged" (Dupré 1981, 82). My own (naturalistic) stance is that there are multiple *legitimate* levels of classification, so long as classifications individuate *natural similarities* that unify members of a class. Some human science classifications, including more 'social' classifications (e.g., 'introversion,' 'puberty,' 'schizophrenia,' 'psychopathy') individuate classes in terms of a set of natural properties. Other classifications (e.g., 'policemen,' 'liberals,' 'widows') individuate groups of people in terms of purely conventional ('social') properties, i.e., properties that do not necessarily have a natural

content of ampliative inferences are informative insofar as they go beyond the surface properties ("nominal essence") used to identify kinds.

basis, but are convenient for some social purpose. The biological kinds discussed in this paper are examples of *more natural classes* that are *definable* in terms of biological criteria. My account of biological kinds (*pace* Dupré) is 'essentialist' insofar as it demands that *some* of the properties constituting kinds must be *intrinsic* biological properties common to kind members, which underwrite the reliable projectable inferences yielded by kind terms.

While human science classifications individuate biological kinds more or less accurately and biological kinds can be defined on multiple levels (e.g., genetic, physiological, neurobiological), evidence suggests that some human classifications accurately track biological kinds. For example, classifications of trisomy disorders (e.g., Down syndrome, Edwards syndrome, Patau Syndrome) uncontroversially distinguish kinds by specific genetic criteria (e.g., possessing an extra twenty-first or eighteenth chromosome). Other medical conditions (e.g., PKU, breast cancer) are defined by biological criteria on multiple levels (e.g., an inherited lack of phenylalanine hydroxylase production, mutation on the PAH gene, positive estrogen or progesterone receptor status, mutations on the *BRCA1* or *BRCA2* genes). Neurobiological definitions of introversion and extraversion (i.e., introverts present higher baseline levels of cortical arousal in the ascending reticular activating system compared to extraverts) have received (construct) validation since the 1960s (Eysenck 1991). While more controversial, psychiatric research has identified neurobiological biomarkers for both the positive (e.g., excessive dopamine activity in the mesolimbic pathway) and negative (e.g., deficient glutamate activity in the prefrontal cortex due to impaired neuroplasticity) symptoms of schizophrenia; and genetic biomarkers (variants on chromosomes 1 and 6) for subtypes have been identified (Tsou 2021). The reliable and stable projectable predictions yielded by these classifications (e.g., PKU infants can be treated by placing them on a low phenylalanine diet, introverts salivate more in

response to a lemon drop than extraverts, negative symptoms of schizophrenia can be treated with glutamate agonist drugs) provide compelling evidence that these classifications accurately individuate stable biological kinds. It is important to note that the projectable inferences of these classifications are *causal consequences* of the *intrinsic* biological properties (e.g., negative symptoms of schizophrenia are associated with deficient glutamate activity in the prefrontal cortex) shared by members of a biological kind.

In response to skepticism about biologizing human kinds, I have elsewhere defended a philosophical account of biological kinds that can explain the projectability of human science classifications (Tsou 2020, 2022). This account combines insights from Richard Boyd's homeostatic property cluster (HPC) theory of natural kinds (Boyd 1999a) and Michael Devitt's intrinsic essentialism (Devitt 2008, 2010, 2023). Boyd's HPC theory is motivated to explain the projectability of the messier classes studied in the special sciences (e.g., biology, psychology, medicine): HPC kind terms yield projectable inferences because members of a class share a family of properties that cluster in a regular and nonaccidental (lawlike) 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* relevant causal structures—that is, the dispositional properties and mechanisms in (2) that cause properties to

cluster in a regular and nonaccidental way—explains successful projectable inferences. According to this "accommodation thesis," successful inductive inferences and explanatory generalizations that are generated within a paradigm offer (abductive) evidence that the posited kinds are *accurately representing* (i.e., 'accommodating') genuine causal regularities in nature. Boyd (1991) writes: "Kinds useful for induction or explanation must always 'cut the world at its joints' in this sense" (139).

Boyd's accommodation thesis is insufficient to explain the projectability of HPC kinds, and his HPC theory should be constrained by the further ('essentialist') requirement that at least *some* of the properties individuated by a kind term are intrinsic (e.g., biological) properties and mechanisms shared by kind members. In the HPC kinds literature, there is disagreement regarding whether *any* of the properties or mechanisms underwriting HPC kinds need to be intrinsic. Philosophers of biology 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 et al. 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. My own view is that—*in order to produce robust and ampliative projectable inferences*—HPC kind classifications (including classifications of human kinds) must individuate some intrinsic properties or mechanisms.[8]

[8] While Boyd argues that species classifications must individuate some intrinsic properties, he does not apply this requirement to HPC kinds more generally. 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

As suggested by essentialist accounts of natural kinds (see Khalidi 2013, ch. 1), it is the intrinsic ('essential') natural properties shared by kind members that explain the projectability of their classifications. On this issue, it is important to distinguish two distinctive roles that mechanisms play in Boyd's 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, descent from a common ancestor) can account for (1), they are insufficient to address (2) because they are specified *too generally* to yield specific inferences about species members (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) or internal mechanisms (e.g., physiological mechanisms) common to kind members.

If some objects of classification in the human sciences are HPC kinds constituted by intrinsic biological properties and mechanisms, then Hacking is incorrect to argue that—*because of classificatory feedback*—human science classifications are incapable of producing projectable inferences. If a human classification individuates an HPC kind underwritten by stable biological

---

importance of intrinsic properties is that these "(imperfectly) shared properties" explain the projectability of HPC terms (Boyd 1999b, 81–82). 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* stance that the "naturalness" of natural kinds is *whatever* reference to such kinds contributes to the accommodation of classifications to causal structures (Boyd 1999a, 158–9). But if the considerations he raises in response to Millikan are relevant for species (i.e., intrinsic properties are needed to explain the projectability of species classifications), they should apply to HPC kinds more generally. Hence, the "naturalness" of HPC kinds should be identified—as in traditional essentialist accounts of natural kinds—with *intrinsic natural properties*.

properties and mechanisms, then the classification will yield stable and ampliative projectable inferences. Moreover, the presence of classificatory feedback will not necessarily require revisions to the classification.[9] This argument clearly applies to medical classifications (e.g., 'HIV,' 'chlamydia,' 'breast cancer') that can change the experiences and behavior of classified patients. While being diagnosed as HIV positive will inevitably change the experiences and behavior of diagnosed individuals, the ways in which individuals change (e.g., feeling stressed or stigmatized) will not require revisions to the meaning of HIV nor the symptoms used to diagnose it. The stability of this object of classification is explained by the fact that HIV is an HPC kind constituted by a partly intrinsic biological essence (i.e., it is an immunodeficiency virus). Classificatory feedback is *irrelevant* to the stability of the HIV classification and the disease it classifies, although classifications can be revised for reasons other than looping effects (e.g., to accommodate new scientific findings, the formulation of new classification systems). It is important to notice that projectable inferences yielded by medical classifications such as HIV—including inferences about treatment (e.g., antiretroviral drug treatment)—are grounded in the fact that the classification refers to a stable biological kind and that patients diagnosed with HIV share relevant biological properties (e.g., chronic immune activation).

The argument above about medical classifications also applies (with qualification) to *more socially constituted* classifications, such as psychiatric classifications (e.g., 'depression,' 'schizophrenia,' 'bipolar disorder') that individuate biological kinds. While being diagnosed with depression will invariably change the experiences and behavior classified individuals, the

[9] In the Hacking–Boyd debate (Hacking 1991a, 1991b; Boyd 1991), Boyd is optimistic that the HPC theory can apply to (projectable) human kind classifications, while Hacking (1991b) is pessimistic, arguing that human kind classifications (e.g., 'multiple personality,' 'homosexuality') are "made up" insofar as "they loop back, interact with, and alter the individuals and the types of behaviour to which they apply" (153). On this issue, I side with Boyd, but argue that a more restrictive (i.e., essentialist) account of HPC kinds is needed to support his conclusion.

specific ways that classified individuals change will not necessarily require revising the *meaning of major depression* nor the primary criteria used to define it (e.g., persistent feelings of sadness). The stability of the classification and projectable inferences associated with the classification (e.g., severe depressive symptoms can be treated with serotonin agonist drugs or REM sleep deprivation) are grounded in the fact that the classification tracks a stable kind.

While classifications of depression (formulated at a sufficiently general level of description) will yield stable projectable inferences, there is evidence that looping effects can create subclasses (or subtypes) of depression. Cross-cultural research indicates that disorders such as schizophrenia, depression, and anxiety disorders appear in all cultures, but their *typical expression* can vary across cultures (Kleinman 1988, chs. 2–3). The largest cross-cultural study of depression identified a 'common core' of depressive signs observed globally: sadness, joylessness, anxiety, tension, lack of energy, decreased interest and concentration, feelings of inadequacy, and feelings of worthlessness (Sartorius et al. 1983, p. 92). Despite this 'universal' form of depression, the ways in which depression is expressed in different cultures varies widely.

For example, in non-Western societies (e.g., China), depression is more typically expressed as somatic complaints (e.g., headaches, dizziness, lack of energy), while in Western societies, depression is more typically expressed as feelings of guilt (Kleinman 1988).[10] More generally, research on 'culture-bound syndromes' (e.g., *susto*, *latah*, *amok*) that only appear in specific cultures indicates that there is a wide variety of *culturally-specific expressions of depression*, *anxiety, or a combination of both*.[11] One way to understand these findings is to

[10] Research suggests that the somatic-depression observed in China is more common globally, while guilt-depression is a cultural variant specific to Western cultures (Kirmayer 2001).

[11] I assume that culture-bound syndromes (CBS) are not different in kind from cultural variants of universal disorders (e.g., guilt-depression) or transient mental disorders (e.g., hysteria). With respect to the *naturalness* of CBS, these disorders are natural to the extent that they are constituted by stable biological mechanisms, which provide the basis for robust projectable inferences. By contrast, the specific expression of such disorders that are

regard *the uniformity of a human kind across* cultures *as a measure of the extent that it is determined by biological properties* (Tsou 2007 2013 2021). This interpretation is consistent with the finding (Marsella 1988) that the symptoms of the most severe and debilitating mental disorders (e.g., schizophrenia) are expressed *more uniformly across cultures* compared to other disorders (e.g., acute depression and anxiety). On this view, some human science classifications (e.g., schizophrenia, bipolar disorder) are 'more natural' than others (e.g., acute depression) because their characteristic signs are *more directly determined by biological mechanisms* and *expressed more uniformly across cultures*. By contrast, human classifications whose characteristic signs are more strongly mediated by social mechanisms (e.g., imitation of stereotypes, role adoption) are 'more social' and more liable to looping effects, which play a central role in the social construction of some human kinds.[12]

The arguments in this section suggest that biological kinds are *stable* targets of human science classifications *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 less prevalent than Hacking's analysis suggests, and they involve classifications that primarily track social properties, rather than biological ones.[13] Conversely, there is evidence that more socially-constituted human science classifications (e.g., depression) that track biological kinds may generate looping effects that result in the creation and stabilization of (cultural) subtypes (e.g., Western guilt-depression, culture bound syndromes).[14] Similar social mechanisms (e.g., imitation

determined by social mechanisms represent socially-constituted aspects of CBS. For a more comprehensive discussion, see Cooper (2010) and Tsou (2021).

[12] For a more comprehensive discussion of the social mechanisms—besides looping effects—involved in socially constructed human categories, see Haslanger (2012), Mallon (2016), and Ásta (2018).

[13] I have elsewhere argued—in response to Mallon's naturalistic theory of social kinds (Mallon 2016)—that human classifications of *purely social kinds* will fail to yield robust and ampliative projectable inferences (Tsou 2020).

[14] It is plausible that some *medical classifications* (e.g., 'obesity') are examples of such socially-constituted classifications of biological kinds (Ritenbaugh 1982).

of stereotypes, role adoption) are involved in 'transient mental disorders' (e.g., 'dissociative fugue,' 'hysteria') that only appear in certain historical times and places (Hacking 1998). This framework provides an elaboration and more precise formulation of Hacking's early distinction between *social* and *natural* human classifications (Hacking 1995b) and his later account of 'interactive and indifferent kinds' (Hacking 1999).[15] On this view, the scope of projectable kinds, which are applicable across different cultures and historical eras, is much broader than the list of basic human natural kinds (e.g., 'mass,' 'longevity') provided by Hacking (1995b, 353). Significantly, projectable kinds will include some more socially-constituted categories (e.g., 'introversion,' 'depression,' 'psychopathy'), so long as the formulated classification accurately tracks a biological kind.

5. Looping Effects Cannot Distinguish the Human from Natural Sciences

Another motivation for Hacking's account of looping effects is to propose a novel way of distinguishing the human from natural sciences. For Hacking, the human and natural sciences can be distinguished in terms of their respective classifications. Whereas classifications in the natural sciences are indifferent kinds that do not produce looping effects, most classifications in

[15] Hacking's account of 'interactive and indifferent kinds' is intended to account for psychiatric disorders (e.g., schizophrenia, autism) that possess both a biological and social basis (Hacking 1999, 108-124). His account suggests that classifications of interactive and indifferent kinds refer to a biological kind ('pathology P'), but these classifications are subject to looping effects, which can change the stereotypical meaning of classifications. Hacking (1999, 119–124) also presents interactive and indifferent kinds in the language of Kripke-Putnam semantics for natural kinds (Kripke 1980, Putnam 1975). Hacking endorses Putnam's idea that the meaning of a kind term includes, both its referent (or 'extension') and its stereotype (Putnam 1975, 245–253). Hacking states that the referent of a classification such as 'autism' is pathology *P*, while its stereotype is the constellation of ideas (including prototypical examples, theories, and attitudes) associated with the classification. In reconciling the apparent dilemma of how a classification can be both indifferent and interactive, Hacking argues that the referent of 'autism' is an indifferent kind, whereas its stereotype is an interactive kind. For criticism of Hacking's 'semantic resolution,' see Murphy (2001) and Tsou (2007).

the human sciences are interactive kinds that produce looping effects (Hacking 1999, 108). Hacking (1995b) writes:

> Responses of people to attempts to be understood or altered are different from the responses of things. This trite fact is at the core of one difference between the natural and human sciences, and it works at the level of kinds. There is a looping or feedback effect involving the introduction of classifications of people. New sorting and theorizing induces changes in self-conception and in behaviour of the people classified. Those changes demand revisions of the classification and theories, the causal connections, and the expectations. Kinds are modified, revised classifications are formed, and the classified change again, loop upon loop. (370)

Hacking's 'trite fact' that classified people will change in response to how they are classified (i.e., classificatory feedback) does not imply that human science classifications must be revised to accommodate those changes (i.e., the looping effect). More generally, this trite fact cannot be used to draw a systematic general distinction between the classifications in the human and natural sciences.

In sections 3 and 4 of this paper, I suggest one line of argument for resisting Hacking's argument that the classifications of human sciences and natural sciences can be distinguished by the presence or absence of looping effects. Some human science classifications are associated with looping effects (e.g., 'hysteria,' 'multiple personality'), but others (e.g., 'schizophrenia,' 'bipolar disorder') are not. More specifically, human science classifications that refer to biological kinds (i.e., HPC kinds constituted by intrinsic biological properties and mechanisms) will produce classificatory feedback, but not necessarily produce looping effects.

In his various accounts of looping effects, Hacking conflates classificatory feedback with looping effects (section 3) to reach the conclusions that human science classifications invariably generate looping effects and cannot be modeled on natural science classifications. Human science classifications that individuate biological kinds (e.g., 'introversion,' 'HIV,' 'hypertension'), while they can affect the people that they classify, do not always generate looping effects in the sense of requiring classificatory revisions to accommodate *the specific ways that the classification changes people*. Hence, some human classifications can be modeled on the natural sciences insofar as they classify stable natural (biological) kinds and yield robust and ampliative projectable inferences.

Another line of argument for resisting Hacking's argument suggests that the manner in which some *natural science classifications* (e.g., 'marijuana,' 'food,' 'bacteria') are formulated (i.e., the *meaning* of these terms) can cause changes to the objects of those classifications, which can sometimes require revisions in the original classification (Bogen 1988, Cooper 2004, Khalidi, 2010; Haslanger, 2012). For example, Bogen (1988) and Cooper (2004) argue that the classification of 'marijuana' *as an illegal drug* (as opposed to a plant found in the wild) in the twentieth century caused a number of changes to marijuana plants, e.g., plants were grown in covert grow operations, the phenotypic traits changed drastically, marijuana strains with much higher THC levels were created.

The feedback generated by classifying marijuana as an illegal drug is analogous to classificatory feedback: the *meaning* of a natural science classification (e.g., illegal drug versus plant) causes changes to its object of classification. A significant point that is not raised in Bogen's and Cooper's analyses is that an analogue of looping effects occurs in the marijuana case: the changes to morphology, anatomy, chemistry, and physiology of marijuana that resulted

from classifying it as an illegal drug has required corresponding taxonomic revisions to the marijuana (*Cannabis sativa*) classification in botany (Small 2015). While most taxonomists recognize a single species of *Cannabis* (*C. sativa*) and regard the many observed subspecies or varieties of marijuana (e.g., narcotic and non-narcotic subspecies) to reflect human domestication practices (Small 2015, pp. 275-6), other taxonomists (e.g., Hillig 2005) argue that there are multiple species (*C. sativa, C. indica, C. ruderalis*) of *Cannabis*. For the purposes of this paper, the important point is that classification of marijuana *as an illegal* drug caused the creation of different *kinds of marijuana* that needed to be accommodated in taxonomic systems. With the partial decriminalization of marijuana in many parts of the world, we can expect more varieties of marijuana to be created (e.g., for recreational use, medicinal use), which will require corresponding revisions to the *Cannabis* classification. Here, we have an example of the *looping effects of natural kinds*: the meaning of some natural science classifications can cause changes to their objects of classification, which require corresponding revisions to the classification.[16]

One could object to the argument above by insisting that the *precise mechanism* by which natural science classifications loop back on themselves is different from looping effects insofar as Hacking's account requires classifications to change their objects of classification (kinds of people) *as a result of classified people being aware of and reacting to how they are classified*. While this objection has some *prima facie* force, there are reasons for rejecting it. First, Hacking (1999) himself has conceded that looping effects do not just include *direct effects* of classifications on objects of classification, but indirect effects. Hacking (1999) writes:

> Autism may seem problematic for my idea of an interactive kind. Autistic children …

[16] Khalidi (2010) offers a similar analysis of looping effects in the classification of wolves and dogs. The classification of wolves (*Canis lupus*) *as domesticated animals* led—through selective breeding—to the creation of dogs and a new species classification (*Canis familiaris*), which required revisions to the original classification, i.e., the addition of *Canis lupus familiaris* as a subspecies.

> have severe problems with communication. So how can the classification interact with the children? Part of the answer is that they are in their own ways aware, conscious, reflective, and, in the experience of those who work with autistic children, very good at manipulating other people. . . . But the example brings out that by interaction I do not mean only the self-conscious reaction of a single individual to how she is classified. *I mean the consequences of being so classified for the whole class of individuals and other people with whom they are intimately connected*. (115, emphasis added)

The last sentence of this passage suggests that classificatory feedback is not limited to *direct effects* of classifications on the expectations and behaviour of *those classified*, but it extends to *indirect effects* of classifications on the expectations and behaviour of *those who interact with classified individuals*.

Hacking never expands on this issue, but if classificatory feedback includes *indirect effects* (as Hacking permits in the autism example), then the indirect effects discussed in the marijuana example and other examples of artificial selection should be regarded as looping effects generated by natural science classifications. Second, given that Hacking is proposing looping effects as marking a fundamental divide between the inherently unstable objects of classification in the human sciences and the stable objects of classification in natural sciences (section 2), his proposed demarcation criterion for the human and natural sciences ought to be evaluated in terms of how successfully looping effects demonstrate this difference. The relevant issue is not whether natural science classifications change its objects of classification *via the specific mechanism of objects being aware of and reacting to how they are classified*, but whether the feedback generated natural science classifications can cause comparable

metaphysical changes in their classificatory objects. Rachel Cooper expresses this argument as follows:

> While it is true that only human kinds are affected by the subjects' ideas, it is also true that only bacteria are affected by antibiotics, and that only domestic animals can be selectively bred. . . . The fact that only human kinds are affected by the subjects' ideas will only be a reason for thinking that human kinds are distinct from natural kinds if extra premise is added, to the effect that being affected by ideas is of greater metaphysical significance than being affected by, say, antibiotics" (Cooper 2004, 79).

By this standard (cf. Allen 2021), Hacking has not shown that the looping effects observed in the human sciences are of greater metaphysical significance than the feedback effects observed for some natural science classifications (e.g., 'bacteria,' 'corn,' 'cow'). Moreover, if one wishes to defend Hacking by appeal to specific mechanisms (e.g., awareness, reaction to being classified) involved in looping effects, it would be more parsimonious to demarcate the human sciences from natural sciences by appeal to their *different objects of study*: objects of human science are aware of and can change in response to how they are classified, while objects of natural science are unaware of how they are being classified. While this position would avoid counterexamples, it reduces to the obvious platitude (Hacking's 'trite fact') that *human sciences study human beings* who are aware of how they are being classified.

The arguments discussed above provide compelling reasons for rejecting Hacking's proposal that the presence or absence of looping effects can serve to demarcate the classifications in the human sciences and the natural sciences, respectively. Some classifications in the human science (viz., those that individuate biological kinds) will generate classificatory feedback, without necessarily generating looping effects. This suggests that *some classifications in the*

human sciences (i.e., classifications of biological kinds) *will not generate looping effects*. Conversely, some classifications in the natural sciences will generate an analogue to looping effects insofar as the meaning of a natural science classification can cause changes to the objects that they classify, which will require corresponding revisions to the classification. This suggests that *some natural science classifications* (e.g., classifications that track artificial selection practices) *generate looping effects* and may even render their objects of classification 'moving targets.'

6. Conclusion

Hacking's account of looping effects accurately describes a phenomenon in the human sciences, wherein the meaning of a classification changes the experiences and behavior of classified people, which require corresponding revisions to the classification. In this article, I objected to three generalizations that Hacking has articulated as consequences of looping effects: (1) looping effects render the objects of classification in the human sciences moving targets, (2) looping effects preclude human science classifications from generating stable and ampliative projectable inferences, and (3) looping effects can distinguish the classifications of the human sciences from those in the natural sciences. Against (1) and (2), I argued that Hacking's generalizations are non-sequiturs that conflate classificatory feedback (i.e., human science classifications change classified people) with looping effects (i.e., human science classifications must be revised to accommodate ways that classified people change). While classificatory feedback is a ubiquitous feature of human science classifications, looping effects are not. In defense of the biologized human kinds eschewed by Hacking, I argued that if a human science classification (e.g., 'HIV,'

‘schizophrenia’) individuates a *biological kind* (i.e., an HPC kind constituted by intrinsic biological properties), then the classification can yield stable and ampliative projectable inferences *in spite of classificatory feedback*. Against (3), I argued that looping is a problematic demarcation criterion because not all human science classifications (e.g., classifications of biological kinds) generate looping effects and some natural science classifications (e.g., ‘bacteria,’ ‘marijuana’) generate looping effects.

In my view, Hacking’s hasty generalizations about interactive kinds, looping effects, and moving targets advanced in his later works (Hacking 1999, 2002b, 2007) betray his *particularist* methodological approach to philosophy (Madsen, Servan, and Øyen 2013, Tsou 2024). Hacking’s philosophical methodology recommends drawing *particular* philosophical conclusions from historical cases and avoiding abstract (‘universal’) generalizations about science. This methodological approach is prominently featured in Hacking’s accounts of entity realism (Hacking 1983) and scientific objectivity (Hacking 2015), and it is encapsulated by his injunction: “let’s get down to work on cases, not generalities” (Hacking 2015, 29). While Hacking (1986: 114) emphasized that there is no single or unified account of making up people, he did not shy away from offering generalizations in later works (Hacking 1998, 1999, 2007). The invalid generalizations that Hacking drew from his analysis of looping effects rested fundamentally on a failure to recognize the distinction between classificatory feedback and looping effects. Given the mass abundance of detailed historical examples of human science classifications (e.g., ‘child abuse,’ ‘multiple personality,’ ‘dissociative fugue,’ ‘autism,’ ‘homosexuality,’ ‘obesity’) that Hacking presented, he can hardly be faulted for failing to examine a sufficient number of cases. As suggested in this article, however, examining a wider range of examples (e.g., examples of human science classifications that fall closer to the ‘natural’

side of the natural-artificial continuum) provide reasons for rejecting his generalizations about the instability of human science classifications, the projectability of human science classifications, and looping effects as a demarcation criterion for distinguishing the human from natural sciences.[17]

[17] I am grateful to Matteo Vagelli, Paul Roth, Janette Dinishak, Mark Risjord, Anya Plutynski, Lydia Patton, Natalia Washington, Joseph McCaffrey, Şerife Tekin, Jamie Shaw, and Sophie Reyes Tsou for helpful comments and criticism. Earlier drafts of this paper were presented at the Philosophy of Ian Hacking Conference at the Hungarian Academy of the Sciences in March 2019, the Philosophy of Medicine and Mental Health Conference at the University of Nebraska at Omaha in April 2019, the Perspectives on the Philosophy of Ian Hacking International Workshop at Ca' Foscari University of Venice in November 2023, and the 5th Undergraduate Philosophy Conference at the University of Texas at Dallas in April 2024. I thank the participants of these conferences for feedback and discussion. I must also express my deep gratitude to Ian Hacking for extended discussion on these issues, when he was an external supervisor for my doctoral dissertation (*The Reality and Classification of Mental Disorders*) at the University of Chicago (c. 2005-2008).

REFERENCES

Allen, Sophie R. 2021. "Kinds Behaving Badly: Intentional Action and Interactive Kinds," *Synthese* 198(Supp. 12): S2927–56.

Anscombe, G. E. M. 1957. *Intention*. Cambridge, MA: Harvard University Press.

Anscombe, G. E. M. 1979. "Under a Description," *Noûs,* 13(2): 219–233.

Ásta 2018. *Categories We Live By: The Construction of Sex, Gender, Race, and Other Social Categories*. New York: Oxford University Press,

Bogen, James 1988. "Comments on 'The Sociology of Knowledge about Child Abuse," *Noûs* 22(1): 65–66.

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

Cooper, Rachel 2004. "Why Hacking is Wrong about Human Kinds," *British Journal for the Philosophy of Science* 55(1): 73–85.

Cooper, Rachel 2010. "Are Culture-Bound Syndromes as Real as Universally-Occurring Disorders?" *Studies in History and Philosophy of Biological and Biomedical Sciences* 41(4), 325–32.

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.

Devitt, Michael 2023. *Biological Essentialism*. Oxford: Oxford University Press.

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

Eysenck, Hans J. 1991. "Dimensions of Personality: The Biosocial Approach to Personality," in Jan Strelau and Alois Angleitner (eds.), *Explorations in Temperament* (pp. 87-103). New York: Springer.

Fagerberg, Harriet 2022. "Reactive Natural Kinds and Varieties of Dependence," *European Journal for Philosophy of Science* 12: 72. https://doi.org/10.1007/s13194-022-00500-x

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 Robert A. Wilson (ed.), *Species: New Interdisciplinary Essays*. Cambridge, MA: MIT Press, 209–28.

Hacking, Ian 1983. *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. Cambridge: Cambridge University Press

Hacking, Ian 1984. "Five Parables," reprinted in Hacking (2002a, 27-50).

Hacking, Ian 1986. "Making Up People," reprinted in Hacking (2002a, 99-114)

Hacking, Ian 1991a. "A Tradition of Natural Kinds," *Philosophical Studies* 61(1-2): 109-126.

Hacking, Ian 1991b. "On Boyd," *Philosophical Studies* 61(1-2): 149-154.

Hacking, Ian 1995a. *Rewriting the Soul: Multiple Personality and the Sciences of Memory*.

Princeton, NJ: Princeton University Press.
Hacking, Ian 1995b. “The Looping Effects of Human Kinds,” in Dan Sperber, David Premack and Anne J. Premack, eds., *Causal Cognition: A Multidisciplinary Debate*. Oxford: Clarendon Press, 351–83.
Hacking, Ian 1998. *Mad Travelers: Reflections on the Reality of Transient Mental Illness*. Charlottesville: University Press of Virginia.
Hacking, Ian 1999. *The Social Construction of What?* Cambridge, MA: Harvard University Press.
Hacking, Ian 2002a. *Historical Ontology*. Cambridge, MA: Harvard University Press.
Hacking, Ian 2002b. “How ‘Natural’ are ‘Kinds’ of Sexual Orientation?” *Law and Philosophy* 21(3): 335-47.
Hacking, Ian 2007. “Kinds of People: Moving Targets,” *Proceedings of the British Academy* 151: 285–318.
Hacking, Ian 2015. “Let’s Not Talk about Objectivity,” in Flavia Padovani, Alan Richardson, and Jonathan Y. Tsou (eds.), *Objectivity in Science: New Perspectives from Science and Technology Studies*, Boston Studies in the Philosophy and History of Science, vol. 310. Cham: Springer, 19-33.
Haslanger, Sally 2012. *Resisting Reality: Social Construction and Social Critique*. Oxford: Oxford University Press.
Hillig, Karl W. 2005. “Genetic Evidence for Speciation in *Cannabis* (Cannabacae),” *Genetic Resources and Crop Evaluation* 52(2): 161-180
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.
Kirmayer, Laurence J. 2001. “Cultural Variations in the Clinical Presentation in Depression and Anxiety: Implications for Diagnosis and Treatment. *Journal of Clinical Psychiatry* 13(Suppl. 13), 22–8.
Kleinman, Arthur 1988. *Rethinking Psychiatry: From Cultural Category to Personal Experience.* New York: Free Press.
Kripke, Saul. 1980. *Naming and Necessity.* Cambridge, MA: Harvard University Press.
Madsen, Ole Jacob, Servan, Johannes, and Øyen, Simen Andersen 2013. “‘I am a Philosopher of the Particular Case’: An Interview with the 2009 Holberg Prizewinner Ian Hacking,” *History of the Human Sciences* 26(3): 32–51.
Mallon, Ron 2016. *The Construction of Human Kinds*. Oxford: Oxford University Press.
Marsella, Anthony J. (1988). “Cross-Cultural Research on Severe Mental Disorders: Issues and Findings,” *Acta Psychiatrica Scandinavica* 78(S344), 7–22.
Millikan, Ruth G. 1999. “Historical Kinds and the ‘Special Sciences,’” *Philosophical Studies* 95 (1–2):45–65.
Marchionni, Caterina, Zahle, Julie, and Godman, Marion 2024. “Reactivity in the Human Sciences.” *European Journal for Philosophy of Science* 14:8. https://doi.org/10.1007/s13194-024-00571-y
Murphy, Dominic 2001. “Hacking’s Reconciliation: Putting the Biological and Sociological Together in the Explanation of Mental Illness,” *Philosophy of the Social Sciences* 31(2): 139–62.
Ritenbaugh, Cheryl (1982). “Obesity as a Culture-Bound Syndrome,” *Culture, Medicine and*

Psychiatry 6(4): 347-61
Rundhardt, Rosa W. 2021. "Reactivity in measuring depression," *European Journal for Philosophy of Science* 11: 77. https://doi.org/10.1007/s13194-021-00395-0
Rundhardt, Rosa W. 2022. "Legitimate Reactivity in Measuring Social Phenomena: Race and the Census," *Philosophy of the Social Sciences* 53(2): 122–41.
Putnam, Hilary. 1975. "The Meaning of 'Meaning'," in Hilary Putnam, *Mind, Language and Reality: Philosophical Papers*, vol. 2. Cambridge: Cambridge University Press, 70-84.
Quine, W. V. 1969. *Ontological Relativity and Other Essays*. New York: Columbia University Press.
Salmon, Wesley C. 1967. *The Foundations of Scientific Inference*. Pittsburgh, PA: University of Pittsburgh Press.
Sartorius, N., Davidian, H., Ernberg, G., et al. 1983. *Depressive Disorders in Different Cultures: Report on the WHO Collaborative Study on Standardized Assessment of Depressive Disorders*. Geneva: World Health Organization.
Small, Ernest. 2015. "Evolution and Classification of *Cannabis sativa* (Marijuana, Hemp) in Relation to Human Utilization," Botanic Review 81(3): 189-294.
Scheff, Thomas J. 1963. "The Role of the Mentally Ill and the Dynamics of Disorder: A Research Framework. *Sociometry* 26(4): 436–53.
Szasz, Thomas S. 1960. "The Myth of Mental Illness," *American Psychologist* 15(2): 113–8
Tekin, Şerife 2011. "Self-Concept through the Diagnostic Looking Glass: Narratives and Mental Disorder," *Philosophical Psychology* 24(3), 357–80.
Tekin, Şerife 2014. "The Missing Self in Hacking's Looping Effects," in Harold Kincaid and Jacqueline A. Sullivan, eds., *Classifying Psychopathology: Mental Kinds and Natural Kinds*. Cambridge, MA: MIT Press, 227–56.
Tsou, Jonathan Y. 2007. "Hacking on the Looping Effects of Psychiatric Classifications: What Is an Interactive and Indifferent Kind?" *International Studies in the Philosophy of Science* 21(3): 329–44.
Tsou, Jonathan Y. 2013. "Depression and Suicide are Natural Kinds: Implications for Physician-Assisted Suicide," *International Journal of Law and Psychiatry* 36(5–6): 461–70.
Tsou, Jonathan Y. 2016. "Natural Kinds, Psychiatric Classification and the History of the *DSM*," *History of Psychiatry* 27(4): 406–24.
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.
Tsou, Jonathan Y. 2022. "Biological Essentialism, Projectable Human Kinds, and Psychiatric Classification," *Philosophy of Science* 89(5): 1155–65.
Tsou, Jonathan Y. 2024. "The Ambiguous Legacy of Kuhn's *Structure* for Normative Philosophy of Science," in K. Brad Wray, ed., *Kuhn's The Structure of Scientific Revolutions at 60.* Cambridge: Cambridge University Press, pp. 217–34.
Wilson, Robert A. 1999. "Realism, Essence, and Kind: Resuscitating Species Essentialism?" in Robert A. Wilson (ed.), *Species: New Interdisciplinary Essays*, Cambridge, MA: MIT Press, 187–207.
Wilson, Robert A., Matthew J. Barker, and Ingo Brigandt 2007. "When Traditional Essentialism Fails: Biological Natural Kinds," *Philosophical Topics* 35 (1–2): 189–215.