

Hacking on the Looping Effects of Psychiatric Classifications: What Is an Interactive and Indifferent Kind?

Jonathan Y. Tsou

This paper examines Ian Hacking's analysis of the looping effects of psychiatric classifications, focusing on his recent account of interactive and indifferent kinds. After explicating Hacking's distinction between 'interactive kinds' (human kinds) and 'indifferent kinds' (natural kinds), I argue that Hacking cannot claim that there are 'interactive and indifferent kinds,' given the way that he introduces the interactive-indifferent distinction. Hacking is also ambiguous on whether his notion of interactive and indifferent kinds is supposed to offer an account of classifications or objects of classification. I argue that these conceptual difficulties show that Hacking's account of interactive and indifferent kinds cannot be based on—and should be clearly separated from—his distinction between interactive kinds and indifferent kinds. In clarifying Hacking's account, I argue that interactive and indifferent kinds should be regarded as objects of classification (i.e., kinds of people) that can be identified with reference to a law-like biological regularity and are aware of how they are classified. Schizophrenia and depression are discussed as examples. I subsequently offer reasons for resisting Hacking's claim that the objects of classification in the human sciences—as a result of looping effects—are 'moving targets'.

1. Introduction

Ian Hacking (1986; 1995a; 1995b, ch. 2; 1999, ch. 4) has described a phenomenon that he calls the 'looping effects of human kinds'. Looping refers to a feedback effect wherein the meaning of a human science classification (e.g., 'schizophrenia', 'multiple personality', or 'depression') affects the behaviour of those who fall under that classification (e.g., an individual diagnosed with depression acts in accordance with the expectations

Jonathan Y. Tsou is a Ph.D. candidate in the Committee on Conceptual and Historical Studies of Science, University of Chicago.

Correspondence to: Committee on Conceptual and Historical Studies of Science, University of Chicago, Room 205, 1126 East 59th Street, Chicago, IL 60637, USA. E-mail: jtsou@uchicago.edu

fostered by that classification). Hacking maintains that looping effects are specific to human science classifications, and much of his discussion has been articulated with reference to psychiatry. Although Hacking is not original in advancing this thesis concerning psychiatric classifications (e.g., see Laing 1967; Scheff 1963; Szasz 1974),¹ he has provided the clearest and most comprehensive discussion of this topic in the context of contemporary philosophy of science.

This paper examines Hacking's analysis of psychiatric classifications, and in particular, his account of *interactive and indifferent kinds*. The paper proceeds as follows. I begin by explicating Hacking's general account of looping, focusing on the account of human kinds presented in *The Social Construction of What?* (Hacking 1999, ch. 4), which invokes a distinction between interactive kinds and indifferent kinds. Subsequently, I examine Hacking's notion of interactive and indifferent kinds, arguing that his presentation of this notion is inconsistent with the way that he introduces the interactive-indifferent distinction. I locate this inconsistency in Hacking's equivocation on the term 'indifferent' in his interactive-indifferent distinction (wherein 'indifference' refers to classifications without looping effects) and his notion of interactive and indifferent kinds (wherein 'indifference' refers to the presence of a stereotypical biological abnormality). Moreover, Hacking's equivocation highlights an ambiguity concerning whether his account of interactive and indifferent kinds is intended to refer to *classifications* or *objects of classification* in the human sciences. I argue that these conceptual problems show that Hacking's notion of interactive and indifferent kinds cannot be based on—and must be clearly separated from—his distinction between interactive kinds and indifferent kinds. In attempting to clarify Hacking's account, I argue that interactive and indifferent kinds should be regarded as objects of classification, presenting schizophrenia and depression as concrete examples of such kinds. On the basis of this revised understanding of interactive and indifferent kinds, I offer some reasons for resisting Hacking's suggestion that the objects of classification in the human sciences are 'moving targets' (i.e., unstable objects of knowledge).

2. Looping Effects, Interactive Kinds, and Indifferent Kinds

In the context of a historical discussion of multiple personality disorder, Hacking describes the looping effects of human kinds as follows:

We tend to behave in ways that are expected of us, especially by authority figures—doctors, for example. Some physicians had [individuals with multiple personalities] among their patients in the 1840s, but their picture of the disorder was very different from the one that is common in the 1990s. The doctors' vision was different because ... the doctors' expectations were different. That is an example of a very general phenomenon: the looping effect of human kinds. 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. Multiple personality is an almost too perfect illustration of this feedback effect. (Hacking 1995b, 21, emphasis added)

Hacking's account of looping concerns the feedback effects surrounding human science classifications, and in particular, how *expectations* engendered by certain classifications

open new possibilities concerning the *kinds of people* that can exist (Hacking 1986). As indicated in the passage above, Hacking believes that looping effects contribute to the evolution of kinds of people, as well as the evolution of scientific classifications. In the context of ‘child abuse’, Hacking writes: ‘The concept of child abuse may itself be so made and molded by attempts at knowledge and intervention, and social reaction to these studies, *that there is no stable object*, child abuse, to have knowledge about’ (Hacking 1995b, 61, emphasis added).

According to Hacking, the human sciences are distinctive insofar as its classifications, ‘human kinds’, will typically have looping effects, whereas the classifications of the natural sciences, ‘natural kinds’, will not. In *The Social Construction of What?*, Hacking (1999 103–108) articulates this position through a distinction between ‘interactive kinds’ (human kinds) and ‘indifferent kinds’ (natural kinds). According to this distinction, *interactive kinds* (e.g., ‘child abuse’, ‘homosexuality’, ‘depression’) are classifications that have looping effects, and *interact* with what they classify. Hacking writes,

‘Interactive’ is a new concept that applies not to people but classifications ... that can influence what is classified ... We are especially concerned with classifications that, when known by people or those around them, and put to work in institutions, change the ways in which individuals experience themselves—and may even lead people to evolve their feelings and behavior in part because they are so classified. (Hacking 1999, 103–104)

In contrast to interactive kinds, *indifferent kinds* (e.g., ‘water’, ‘sulphur’, ‘lemon’) are classifications that do not have looping effects, and do not interact with what they classify. On the basis of this distinction, Hacking (1999, 108) claims that the social sciences can be distinguished from the natural sciences insofar as the classifications of the latter are indifferent kinds, whereas the classifications of the former are (typically) interactive kinds.

While Hacking’s position appears to be amenable to constructionist perspectives on the human sciences, he has made it clear that he has no desire to be a social constructionist, defending a middle ground between realism and social constructionism (what he calls ‘dynamic nominalism’). Hacking writes,

One of the defects of social construction talk is that it suggests a one-way street: society (or some fragment of it) constructs the disorder (and that is a bad thing, because the disorder does not really exist as described, or would not really exist unless so described). By introducing the idea of an interactive kind, I want to make plain that we have a two-way street, or rather a labyrinth of interlocking [causal] alleys. (Hacking 1999, 116)

As indicated here, Hacking’s account of interactive kinds is not meant to imply that such kinds have no basis in reality, but to highlight the complex interactive relations (between reality, classifications, and things classified) specific to such kinds.

For the purposes of this paper, it is important to note that Hacking’s (1999, 103–106) distinction between interactive kinds and indifferent kinds is intended to distinguish between two types of *scientific classifications*, rather than two ‘kinds of things’ in nature. In his seminal paper on looping effects, Hacking emphasized that his account of human

kinds was intended to describe ‘systems of classifications—rather than people and their feelings’ (Hacking 1995a, 352). Generally, Hacking has been consistent in limiting his various accounts of human kinds to *classifications*. I will subsequently suggest, however, that when Hacking introduces his notion of interactive and indifferent kinds, he is not consistent in this regard.

3. Interactive and Indifferent Kinds in Psychiatry

Hacking maintains that some human science classifications (e.g., ‘schizophrenia’, ‘mental retardation’, ‘autism’) are both interactive and indifferent (Hacking 1999, 108–124). In the remainder of this paper, my attention will be focused on Hacking’s account of interactive and indifferent kinds in psychiatry.

Psychiatric conditions such as schizophrenia, autism, or mental retardation raise difficulties for Hacking’s claim that the classifications for these conditions are interactive. Since individuals with such conditions (e.g., a child with a serious form of autism) are *not fully aware* of their condition or how they are classified, Hacking’s contention that the classification ‘autism’ *interacts* with autistic children is problematic. In response to such an objection, Hacking writes,

Autism may seem problematic for my idea of an interactive kind. Autistic children ... 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. (Hacking 1999, 115)

The last sentence of this passage suggests that looping is not limited to *direct effects* on the expectations and behaviour of *those classified*, but also includes *indirect effects* on the expectations and behaviour of *those who interact with classified individuals*.² Thus, Hacking views the looping effects of human kinds to be far-reaching, and his more qualified account of looping indicates the way in which he envisages classifications such as ‘autism’ to be interactive kinds.³

If autism and schizophrenia are interactive in the way suggested above, how can they be indifferent? Hacking’s answer to this question appeals to the biological bases of such conditions. Hacking writes,

There is a deep-seated conviction that retarded children, schizophrenics, and autistic people suffer from ... fundamental neurological or biochemical problems ... No one maintains that mental retardation is a single disorder, but *many believe that specific types of retardation have clear biological causes, to the extent that we can say these disorders simply are biological in nature* ... We need not argue that nearly all children diagnosed with autism today have exactly one and the same biological disorder. We need only hold possible that there are a few (possibly just one) basic fundamental biological disorders that produce the symptoms currently classified as autistic ... Let us posit that there is a pathology *P*, no matter how it will be identified. By hypothesis the

pathology *P* will be an indifferent kind. The neuro-geno-biochemical state *P* is not aware of what we find out. It is not affected simply by the fact that we have found out about it ... In more traditional jargon, *P* would be a natural kind. (Hacking 1999, 116–117, emphasis added)

As indicated here, Hacking identifies the indifferent aspect of interactive and indifferent kinds with identifiable biological regularities that are associated with conditions such as mental retardation, autism, or schizophrenia. In postulating that the symptoms of autism are caused by pathology *P*, Hacking is not suggesting that *all* forms of autism (viz., the pervasive developmental disorders that make up the so-called ‘autism spectrum’) will someday be reducible to pathology *P*, but that a single biological cause may be discovered for a certain sub-type of autism (e.g., ‘autistic disorder’).⁴

Hacking (1999, 119–124) also presents interactive and indifferent kinds in the language of Kripke-Putnam semantics for natural kinds (Kripke 1980; Putnam 1975) as a ‘semantic resolution’ to the dilemma of how something can be both an interactive kind and indifferent kind.⁵ Hacking endorses Putnam’s (1975, 245–253) idea that the meaning of a kind term includes, both its referent (or ‘extension’) and its stereotype. In applying this framework to psychiatric classifications, 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) currently associated with the classification. In reconciling the apparent dilemma of how something can be both an indifferent kind and an interactive kind, Hacking argues that the referent of ‘autism’ is an indifferent kind, whereas its stereotype is an interactive kind.

Hacking’s notion of interactive and indifferent kinds is formulated to accommodate the putative fact that some psychiatric conditions (e.g., schizophrenia and autism) have both a biological and social basis. As indicated above, Hacking also introduces the notion to articulate a conciliatory position on the *reality* of human kinds:

something can apparently be both socially constructed and yet ‘real’... childhood autism *is* (is identical to) a certain biological pathology *P*, and so is a ‘natural’ kind or an indifferent kind. At the same time, we want to say that childhood autism is an interactive kind, interacting with autistic children, evolving and changing as the children change. (Hacking 1999, 119, emphasis in the original)

Hacking suggests that the biological (or indifferent) aspect of interactive and indifferent kinds constitutes their real basis in nature. In this manner, Hacking’s account of interactive and indifferent kinds is able to account for psychiatric conditions associated with predictable biological regularities, and also account for the fact that the classifications of these conditions can have (direct or indirect) feedback effects on those classified.

4. Problems with Hacking’s Presentation of Interactive and Indifferent Kinds

As outlined above, Hacking characterizes an interactive and indifferent kind as a natural kind (viz., a biological pathology) that is subject to looping effects. This formulation, however, is not consistent with the way that Hacking introduces his distinction between interactive kinds and indifferent kinds. In the following section, I argue that Hacking’s distinction between interactive kinds and indifferent kinds—as a

distinction between two types of classifications—renders his notion of interactive and indifferent kinds (which is supposed to be based on this distinction) incoherent. My argument is intended to highlight the inconsistency between Hacking's formulation of his interactive-indifferent distinction, on the one hand, and his notion of interactive and indifferent kinds, on the other.

Hacking (1999, 103–108) *defines* interactive kinds and indifferent kinds in a mutually exclusive manner. Interactive kinds are 'classifications that, when known by people or by those around them ... change the ways in which individuals experience themselves' (Hacking 1999, 104); indifferent kinds are classifications that do not affect what they classify. As Hacking puts it, 'The classification "quark" is indifferent in the sense that calling a quark a quark makes no difference to the quark' (Hacking 1999, 105). As such, Hacking's distinction between interactive kinds and indifferent kinds is *defined solely with reference to looping effects*: Interactive kinds have looping effects, whereas indifferent kinds do not. For this reason, Hacking rejects the idea that the distinction between interactive kinds and indifferent kinds is a distinction of degree (Hacking 1999, 107–108); the manner in which he defines this distinction renders it a distinction of kind.

Given Hacking's manner of defining interactive kinds and indifferent kinds as 'classifications that affect their objects of study' and 'classifications that do not affect their objects of study', respectively, he is not entitled to maintain that a classification such as autism can be *both* interactive and indifferent. Either this classification will have looping effects (interactive) or it will not (indifferent). The fact that autism has a predictable biological basis does not entitle Hacking to say—without equivocation—that autism is an indifferent kind, since Hacking defines indifference in terms of 'lack of looping effects'. Consider two ideas presented by Hacking:

- 1) *The distinction between interactive kinds and indifferent kinds*—Looping is the distinguishing feature that separates interactive kinds from indifferent kinds; interactive kinds have looping effects, while indifferent kinds do not (Hacking 1999, 103–108).
- 2) *Interactive and indifferent kinds*—There are some human kinds that are both indifferent kinds and interactive kinds (Hacking 1999, 115–120).

Put in this way, it is clear that proposition 2 is inconsistent with proposition 1. At the very least, this inconsistency reveals an underlying tension in Hacking's *presentation* of his account of human kinds. Hacking cannot claim—without contradiction—that looping is the sole criterion for distinguishing interactive from indifferent kinds, while *simultaneously* allowing for kinds that are both interactive and indifferent (i.e., interactive and indifferent kinds cannot be 'classifications that both possess and do not possess looping effects'). What this shows definitively is that interactive and indifferent kinds cannot be articulated with reference to Hacking's distinction between interactive kinds and indifferent kinds.

The source of tension in Hacking's presentation of human kinds, I think, is due to his equivocation on the term 'indifferent' in propositions 1 and 2 above. In proposition 1, Hacking understands indifference in terms of lack of looping effects (i.e., in the sense 'that calling a quark a quark makes no difference to the quark'). This is not the same

meaning that he attaches to the term in the context of proposition 2, where indifference refers to an identifiable biological pathology (i.e., ‘pathology *P*’). While Hacking may be technically correct to say that a biological pathology associated with, e.g., schizophrenia, ‘is not affected simply by the fact that we have found out about it’ (Hacking 1999, 117), he conflates lack of looping effects, on the one hand, with the existence of identifiable biological regularities for certain psychiatric conditions, on the other.

A related point to notice is that in the context of his discussion of interactive and indifferent kinds, Hacking is not discussing features of classifications per se, but objects of classification. That is, although Hacking intends to provide an account of classifications in discussing human kinds, he covertly slips into the realm of ‘kinds in nature’ when he raises his notion of interactive and indifferent kinds. Recall Hacking’s discussion of autism as an interactive and indifferent kind: ‘childhood autism *is* (is identical to) a certain biological pathology *P*, and so is a “natural” kind’ (Hacking 1999, 119, emphasis in the original). Since ‘classifications’ are not usually identified with ‘biological pathologies’, I submit that Hacking is inadvertently discussing features of things in nature, and not classifications. Also consider Hacking’s claim that ‘many believe that specific types of retardation have clear biological *causes*, to the extent that we can say these disorders simply are biological’ (Hacking 1999, 116, emphasis added). This indicates quite clearly that Hacking’s account of interactive and indifferent kinds does not concern classifications (or human kind *terms*), but kinds of things in nature. This exposes another inconsistency in Hacking’s treatment of propositions 1 and 2 above. The distinction being drawn in proposition 1 concerns a feature of classifications, whereas proposition 2 concerns a feature of objects of classification. Thus, the tension in Hacking’s overall account also stems from the fact that he uses the term ‘indifferent’ equivocally to refer to a feature of classifications (when he discusses the interactive-indifferent distinction) and of objects of classification (when he discusses the notion of interactive and indifferent kinds).⁶

Hacking’s equivocation in this last regard is also evident in his ‘semantic resolution’ to the dilemma of how something can be both an interactive kind and an indifferent kind. Hacking’s resolution amounts to the claim that the *referent* of a human science classification such as ‘autism’ is the pathology *P* that causes that condition (indifferent kind), while the *stereotype* (i.e., current ideas) associated with the classification will have looping effects (interactive kind). In this formulation, notice that ‘indifferent kind’ does not refer to a *classification* without looping effects, but the *causes* of a condition in nature (pathology *P*). For this reason, Hacking’s semantic resolution fails insofar as it does not genuinely show how ‘autism’ is an indifferent kind in the sense of being a ‘classification without looping effects’ (cf. Murphy 2001, 152–155).

The equivocations that I have brought attention to in this section are meant to show that Hacking’s notion of interactive and indifferent kinds cannot be based upon—and should be clearly separated from—his distinction between interactive kinds (classifications that affect their objects of study) and indifferent kinds (classifications that do not affect their objects of study). While I do not think that these conceptual flaws in Hacking’s presentation are necessarily fatal to his overall theory of human kinds, they do motivate a clarification of what Hacking means by interactive

and indifferent kinds. As suggested above, what Hacking intends to identify with his notion of interactive and indifferent kinds is a certain type of object of classification in nature, rather than a certain type of scientific classification. I articulate this notion of interactive and indifferent kinds further in the following section.

5. Reconsidering Interactive and Indifferent Kinds

In this section, I attempt to address the aforementioned problems in Hacking's presentation, by clarifying and elaborating on his notion of interactive and indifferent kinds.⁷ My analysis will concern interactive and indifferent kinds as kinds of things in nature, i.e., as objects of classification rather than classifications. While Hacking wants to restrict his account of human kinds to classifications—as shown above—his account of interactive and indifferent kinds ultimately concerns objects of classification (or 'kinds of people'). To elaborate the nature of such objects of classification, I clarify what is 'indifferent' and 'interactive' in Hacking's notion of interactive and indifferent kinds. My presentation focuses on schizophrenia and depression, and in particular, the physiological bases for these conditions.

As outlined above, the *indifferent* part of Hacking's notion of interactive and indifferent kinds refers to biological regularities associated with certain stereotyped sets of abnormal behaviours that constitute a particular *kind of person* (e.g., the 'schizophrenic' or the 'manic-depressive'). While Hacking discusses such regularities in general and somewhat vague terms ('pathology *P*'), I think that—for the purposes of clarifying what interactive and indifferent kinds amount to—it is more useful to discuss the details of such regularities. In my opinion, the most convincing arguments for the *reality* of certain conditions come from research on physiological or biochemical pathways associated with some psychiatric conditions.⁸ Below, I discuss some of these findings with reference to research on schizophrenia and depression.

In the psychopathology literature, it is well established that the 'positive symptoms' of schizophrenia (namely, hallucinations and delusions) are correlated with predictable abnormal physiological states. Specifically, it is known that excessive dopamine activity in the mesolimbic pathway (which projects from the midbrain in the ventral tegmentum to the nucleus accumbens) is correlated with positive symptoms (Davison and Neale 1996, ch. 14; Carlson 1999, 440–443; Millon, Blaney, and Davis 1999, ch. 11). This suggests that some forms of schizophrenia (e.g., paranoid schizophrenia) have clear and predictable biological bases. Evidence for this thesis, the 'dopamine hypothesis', is provided by pharmacological research that demonstrates 'reversibility effects'. Not only can the positive symptoms of schizophrenia be reliably alleviated by anti-psychotic drugs, which reduce dopamine activity in the mesolimbic pathway, but in sufficient doses, drugs with opposite pharmacological effects (e.g., cocaine and other stimulant drugs), which increase dopamine activity in the same pathway, can *induce* the positive symptoms of schizophrenia (McKim 2000, ch. 12). This hypothesis is also supported by the facts that anti-psychotic drugs produce side-effects similar to Parkinson's disease, and that Parkinson's is known to be caused, in part, by low levels of dopamine.⁹ Taken together, this indicates convincingly that there is a clear

and predictable biological basis for certain forms of schizophrenia (e.g., paranoid schizophrenia).

While research on depression is more controversial (especially given the higher, and more variable, reported prevalence rates of depression, compared with schizophrenia), research has uncovered similar biological regularities. Specifically, it has been suggested that the feelings of sadness associated with depression are caused by an underactivity of monoamine neurotransmitters, especially serotonin and norepinephrine (Davison and Neale 1996, 241–242; Carlson 1999, 451–454; Millon et al. 1999, ch. 7). Evidence for this thesis, the ‘monoamine hypothesis’, is also provided by pharmacological research that demonstrates reversibility effects. Not only can the symptoms of depression be alleviated by monoamine agonist drugs (e.g., monoamine oxidase inhibitors, tricyclic antidepressants, and Specific Serotonin Reuptake Inhibitors), which function to increase the activity of serotonin and norepinephrine, but drugs with opposite pharmacological effects (e.g., monoamine antagonists such as Reserpine), which decrease the activity of these neurotransmitters, can *induce* the symptoms of depression (Carlson 1999, 452–454).¹⁰ Taken together, this research provides evidence for a predictable biological basis for some forms of depression (e.g., uni-polar depression).

My aim in discussing the physiological bases for conditions such as schizophrenia and depression is to present some concrete and plausible examples of what Hacking identifies as indifferent (‘pathology *P*’) in interactive and indifferent kinds. In the examples presented, what can be identified as ‘indifferent’ are predictable biological pathways and physiological abnormalities associated with schizophrenia and depression. Like Hacking, I take these biological regularities to count as evidence for the reality of these conditions, and indicate the way in which such conditions approach the traditional ideal of ‘natural kinds’.¹¹ As indicated above, the distinguishing feature of such regularities is their law-like character, and *not* the fact that classifications of these conditions will not have looping effects. This also indicates the precise point where Hacking conflates biological regularities with lack of looping effects. While some conditions (e.g., schizophrenia or anxiety disorders) may exhibit biological regularities, this will imply nothing about whether *their classifications* will have looping effects. While Hacking (1999, 117) *identifies* such biological regularities with lack of looping effects, he confuses a feature of *kinds of things in nature* (a law-like character) with a feature of *classifications* (lack of looping effects).

If the ‘indifferent’ aspect of interactive and indifferent kinds refers to biological regularities associated with certain objects of classification, then what does the ‘interactive’ aspect of interactive and indifferent kinds refer to? I would suggest that ‘interactive’—in this context—refers to the *awareness* that some objects of classification (i.e., kinds of people) possess regarding how they are classified. It is worth noting here that one would expect *all human objects of classification* to possess this characteristic, although one would also expect there to be variability on this characteristic depending on *how aware* different kinds of people (e.g., the mentally retarded child versus the depressed adult) are of how they are classified. All things being equal, one would expect kinds of people who possess a greater awareness of their social environment to have a

greater susceptibility of being influenced by how they are classified. For this reason, Hacking's presentation of autism as an interactive and indifferent kind is a somewhat confusing example, given that the way in which autistic children are 'interactive' is an atypical form of awareness (see note 3). While interactive and indifferent kinds may sometimes refer to kinds of people (e.g., autistic children or schizophrenic adults) who are not fully aware of how they are classified, they may more typically refer to kinds of people (e.g., depressed adults or anxious adults) who are directly aware of how they are classified.

These observations motivate a more precise formulation of Hacking's account of interactive and indifferent kinds. According to the analysis here, Hacking's notion of interactive and indifferent kinds refers to certain objects of classification (or kinds of people): 1) that can be identified with reference to a law-like biological regularity (e.g., excessive dopamine activity in the mesolimbic pathway causing the positive symptoms of schizophrenia); and 2) that are aware—in varying degrees—of how they are classified. I have presented schizophrenia and depression as examples of such interactive and indifferent kinds. What this formulation explicitly acknowledges—contra Hacking—is that interactive and indifferent kinds refer to certain kinds of people (rather than a certain kind of classification), whereas the distinction between interactive kinds and indifferent kinds refers to certain kinds of classifications. Accordingly, Hacking's notion of interactive and indifferent kinds should be considered independently of his distinction between interactive kinds and indifferent kinds. This clarification removes the apparent incoherence between Hacking's interactive-indifferent distinction and his notion of interactive and indifferent kinds.

6. The Stability and Instability of Objects of Classification in Psychiatry

At this point, I want to address some related issues concerning the stability of objects of classification in psychiatry, and the human sciences more generally. Hacking suggests that because of looping, objects of classification in the social sciences are unstable or 'on the move'. Hacking writes,

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.* (Hacking 1999, 108, emphasis added)

Hacking's suggestion here is quite radical (also see Hacking, 2007). His claim is that the kinds of people (e.g., autistic children or schizophrenics) classified by interactive kinds ('autism', 'schizophrenia') will change in lieu of looping effects such that there is *no stable object of knowledge to study*. The idea is that because of the constant dynamics between social science classifications and people being classified, the kinds of people being classified are constantly changing in response to how they are classified, and in this precise sense are 'on the move'. Below, I argue that Hacking's claim is a misleading generalization, which requires qualification. I argue that interactive and indifferent kinds—understood as objects of classification—represent a class of objects that are

stable. More specifically, I contend that the indifferent part of interactive and indifferent kinds (i.e., the law-like biological regularity associated with a particular kind of person) can be understood as a *stable object of classification* that does not change as a result of looping effects. As a more general objection, I argue that Hacking's conclusion that classifications with looping effects render their objects of classification unstable ultimately rests on a conflation of—and failure to distinguish between—weak and strong implications of looping effects.

An area of empirical research that is helpful for addressing issues concerning the stability-instability of objects of psychiatric classification is cross-cultural research on psychiatric conditions. Cross-cultural research indicates that conditions such as schizophrenia, depression, and anxiety disorders are ubiquitous across different cultures, but that the characteristic behavioural *expression* of these conditions will (in varying degrees) differ between cultures (see Kleinman 1988, chs. 2–3). To make sense of these findings, I would suggest that the similarities across cultures represent indifferent (biological) aspects of such conditions, while the differences represent interactive (cultural) aspects. Moreover, some conditions (e.g., paranoid schizophrenia) appear more uniformly across cultures compared to other conditions. I would further suggest that the *uniformity of a condition across cultures* can be interpreted as a measure of the extent that a condition is indifferent.¹² While there is a sense in which Hacking is correct to suggest that human kinds are 'on the move', there is an important sense in which some interactive and indifferent kinds (i.e., more indifferent conditions such as paranoid schizophrenia) represent a subset of objects that are *stable*. Since these kinds will be associated with physiological regularities that are uniform across cultures, and assuming that these regularities do not change in lieu of classificatory looping effects,¹³ some interactive and indifferent kinds are not as unstable as Hacking suggests.

The analysis above suggests that there is at least one sense in which Hacking is incorrect to suggest that—because of looping effects—there are no stable objects of classification in the human sciences. A minimal assumption of my analysis is that certain sub-types of a disorder (e.g., paranoid schizophrenia) can legitimately be identified with reference to a clearly specified biological pathology. On this understanding, biological pathologies that cause stereotyped abnormal behaviours (e.g., excessive dopamine activity in the mesolimbic pathway causes hallucinations and delusions) represent one type of object of classification in psychiatry (*viz.*, objects that capture the casual essence of a condition). If one accepts this assumption, then Hacking is incorrect to claim that there are no stable objects of study in psychiatry. To establish this claim, Hacking would need to show that *the typical biological or physiological process that leads to abnormal behaviour is changed because of looping effects*. There is no good evidence for thinking that this is a possible consequence of looping effects.

This last argument can be generalized as an objection that Hacking's argument regarding 'moving targets' fails to distinguish between two different types of implication of looping effects. For simplicity, call these implications:

- (1) *Weak implications of looping*—Individuals' experiences and behaviours are altered in response to looping effects.

- (2) *Stronger implications of looping*—Individuals' experiences and behaviours are altered in response to looping effects *to the extent that the defining criteria for that classification change.*

I take it that implications of type 1 are a ubiquitous—and perhaps trivial—consequence of human science classifications. For the purposes here, it is important to notice that implication 1 does not necessarily render the objects of human science classifications *unstable* in the way suggested by Hacking. While a person who is diagnosed with depression will inevitably be affected by how she is classified, the way in which looping effects change the experiences and behaviour of individuals falling under this classification will not necessarily require a corresponding revision in the classification 'depression' or what it means to be 'clinically depressed'. However, the consequences of looping are sometimes more drastic as specified in implications of type 2. In these cases, looping effects change the experiences and behaviour of individuals—in a uniform manner for a significant number of individuals who fall under that classification—such that the *definitive characteristics* of that classification (i.e., the criteria that constitute membership for a particular classification) must be revised. A good example of implications of type 2 is Hacking's (1995b) historical discussion of multiple personality disorder. When the multiple personality classification was first introduced in the latter part of the 19th century in France, it was regarded as a special case of hysteria. Much later, in the 1970s, a number of case histories were reported in the United States that connected multiple personality with psychological trauma from child abuse (i.e., sexual abuse). This connection between multiple personality and child abuse, Hacking contends, drastically changed the experiences and behaviour of individuals who were classified as multiples. Hacking writes,

The intimate relationship connecting recovery of traumatic memories of child abuse with multiple personality is no accident . . . Dissociation is explained as coping mechanism. The multiple comes to understand that she is as she is now because of the way she deployed coping mechanisms in the past. A narrative structure is available that can then be filled in with the appropriate states. (Hacking 1995b, 256)

What is significant in this story is that by the mid-1980s, the classification 'multiple personality disorder' was transformed such that repeated child abuse was understood to be *the cause* of multiple personality disorder, and regarded as an essential feature of that classification. In this case, looping effects changed the experiences and behaviours of individuals to the extent that the defining characteristics of the classification 'multiple personality' needed to be revised.

The distinction between implications of type 1 and type 2, above, is meant to highlight Hacking's tendency to conflate and obscure a significant difference that can be drawn with respect to the possible consequences of looping effects. While Hacking suggests that looping effects render the objects (or 'targets') of human science classifications unstable (or 'moving targets'), for him to legitimately draw this conclusion, he must show that the consequences of looping are like those specified in implications of type 2. As suggested here, Hacking's analysis fails to establish this point, and obscures the fact that the consequences of looping effects are often more trivial, as specified in

implications of type 1. Hence, Hacking commits a hasty generalization when he argues from the putative existence of looping effects to the instability of objects of classification in the human sciences. While looping effects sometimes imply the instability of objects of classification, this is not always the case. In this section, I have attempted to show that some objects of classification in the human sciences can remain stable *despite looping effects*.

7. Conclusion

In this paper, I have argued that Hacking's presentation of his interactive-indifferent distinction and notion of interactive and indifferent kinds is incoherent. I located this incoherence in Hacking's equivocation on the term 'indifferent', and in his inconsistent application of this term to classifications (in his interactive-indifferent distinction) and to objects of classification (in his notion of interactive and indifferent kinds). In attempting to clarify Hacking's account of human kinds, I have claimed that his account of interactive and indifferent kinds makes most sense as referring to objects of classification or kinds of things in nature. From this perspective, I characterized such kinds as being associated with a predictable biological basis, while possessing some awareness of how they are classified. This account can accommodate conditions such as schizophrenia, depression, and anxiety disorders. However, the classifications for these conditions will be interactive kinds.

On the view suggested in this paper, Hacking's (1999, 108) claim that the objects classified by interactive kinds are 'moving targets' is misleading. If an interactive kind (e.g., the classification 'paranoid schizophrenia') refers to an interactive and indifferent kind (i.e., the object of that classification), Hacking has not provided convincing reasons for believing that the looping effects of this classification will render its object of classification unstable insofar as the biological abnormality that constitutes that classification (viz., excessive dopamine activity in the mesolimbic pathway) will not be altered because of looping effects. More generally, Hacking's analysis implicitly assumes that looping effects will lead to revisions in the criteria that constitute membership for a particular classification. I argued that this implication need not follow from the putative existence of looping effects, and that the implications of looping are often more trivial. At the very least, Hacking needs a more developed argument to provide a convincing case for his contention that the targets of human kind classifications—because of looping effects—are *unstable* objects of knowledge.

Acknowledgements

I owe thanks to Ian Hacking, Robert Richards, Trevor Pearce, Rachel Ponce, Cecelia Watson, Dana Rovang, Anjan Chakravartty, Gregory Radick, Andrew Bailey, Karyn Freedman, Karen Houle, James McAllister, and two anonymous referees of this journal for helpful comments and suggestions. An earlier draft of this paper was presented at the annual meeting of the Canadian Society for the History and Philosophy of Science (CSHPS) at York University in May 2006.

Notes

- [1] In the 1960s and 1970s, anti-psychiatrists and labeling theorists discussed many of the same issues addressed by Hacking's account of looping. Other social scientists (e.g., see Howard 1985) have articulated these issues in terms of the *reflexivity* of human subjects, i.e., human subjects are 'directed or turned back upon themselves' insofar as they are aware of and will react to how they are categorized. Hacking (1986, 2004) has presented his account as revisiting themes from Foucault's (2006) historical studies on madness.
- [2] Several commentators (Bogen 1988; Cooper 2004, 78–80; Schmaus 1992, 169–171) have argued that Hacking fails to demonstrate that human kind classifications differ significantly from natural kind classifications, given that some of the latter (e.g., 'marijuana' or 'dog') are subject to similar feedback effects as the former (cf. Ereshefsky 2004, 913–916). When Hacking allows *indirect* feedback effects (as in the case of autism) to count as looping effects, he must concede that natural kind classifications may possess analogous *indirect* looping effects. In response to these critics, however, Hacking could maintain that since humans kinds are *aware* of the way that they are classified in a way that natural kinds are not, human kind classifications possess another level of feedback (as a result of *direct* looping effects) that are not characteristic of natural kind classifications. This response would save Hacking's contention that human kind terms significantly differ from natural kind terms.
- [3] In section 5 of this article, I suggest that autism, schizophrenia, and mental retardation are not the best examples of interactive and indifferent kinds insofar as such kinds may more typically refer to kinds of people (e.g., depressed adults or anxious adults) who—in contrast with Hacking's examples—are *directly aware* of how they are classified. For the purposes here, it is important to note that interactive and indifferent kinds can be associated with direct feedback effects (in addition to the indirect feedback effects stressed by Hacking).
- [4] Hacking does not discuss in detail any of the contemporary biological research on autistic disorder; research indicates that this particular form of autism is associated with abnormalities in the medial temporal lobe, the brain stem, and the cerebellum (Carlson 1999, ch. 16).
- [5] Hacking's (1999) semantic resolution is also meant to explain how something can be both real and socially constructed (119), and he presents it as 'putting a theory of reference alongside social construction' (122). For a more comprehensive and critical discussion of Hacking's semantic resolution, see Murphy (2001).
- [6] There are other relevant inconsistencies in Hacking's treatment of interactive and indifferent kinds worth mentioning. For example, Hacking (1999, 118–119) clearly has objects of classification (rather than classifications) in mind when he briefly presents interactive and indifferent kinds in diathesis-stress terminology. Hacking suggests that the 'predisposing cause' (diathesis) of a condition can be understood as its indifferent part, while its 'occasioning cause' (stress) can be understood as its interactive part. Here, both the indifferent part (biological cause) and interactive part (social cause) of interactive and indifferent kinds refer to things in nature, rather than classifications. It is also worth mentioning that in this formulation 'interactive' does not even refer to looping effects, since occasioning causes refer to *traumatic life events* (e.g., the death of a spouse), and nothing to do with the looping effects of scientific classifications.
- [7] Hacking has brought my attention to a forthcoming paper (Hacking 2007) that is not liable to the criticism articulated in the previous section. In his forthcoming paper, 'Kinds of People: Moving Targets', Hacking distinguishes five different elements of the human sciences that interact with one another: (a) classifications; (b) people; (c) institutions; (d) knowledge; and (e) experts. While I have criticized Hacking for failing to distinguish between classifications and people (or 'objects of classification'), his forthcoming analysis avoids this difficulty. Accordingly, my criticism of Hacking in the previous section is limited to his presentation in *The Social Construction of What?* (Hacking 1999). Hacking's latest analysis, however, is liable to the criticism (that I raise in section 6 of this paper) that

he has not persuasively shown that objects of knowledge in the human sciences are ‘on the move’ or unstable.

- [8] For a general discussion of pathway research, see Thagard (2003).
- [9] Since the dopamine hypothesis was suggested—mainly on the basis of research on the older ‘typical’ anti-psychotic drugs that work by blocking D₂ receptors—other neurotransmitters (e.g., serotonin) and brain areas have been implicated in schizophrenia (see Carlson 1999, 440–443; McKim 2000, 274–276).
- [10] All of the antidepressant drugs mentioned here are monoamine agonists, which function to increase the activity of the monoamine neurotransmitters: dopamine (DA), norepinephrine (NE), and serotonin (5-HT). Most antidepressants selectively alter NE and 5-HT, while not affecting DA. For a more comprehensive discussion of the various antidepressant drugs and neurobiological theories of depression, see Carlson (1999, 451–454), McKim (2000, 289–292), and Millon et al. (1999, 167–175).
- [11] As an ideal, I assume that natural kinds in psychiatry are discrete, non-arbitrarily bounded classes of abnormal behavior that share a specific etiology (e.g., a genetic cause or a physiological cause); to adopt a phrase from Putnam (1975), natural kinds possess the same ‘general *hidden structure*’ (235, emphasis in the original). I further assume that the mental disorders listed in DSM IV (TR) (*Diagnostic and Statistical Manual*, 4th ed., text revision) (APA 2000) take on a variety of structural forms, and that only a few approach this ideal of natural kinds as a matter of degree (see Haslam 2002). To be clear, my analysis assumes that there are discrete mental disorders that are structurally similar with respect to etiology. I have listed paranoid schizophrenia (and, more tentatively, major depression) as candidates that approach this ideal; other disorders that I would consider are Down syndrome, panic disorder, and bipolar disorder. For a more comprehensive discussion of the issue of mental disorders and natural kinds, see Zachar (2000), Haslam (2002), Cooper (2005, ch. 2), and Murphy (2006, ch. 9). For discussion of Hacking’s own position on natural kinds, see Hacking (1990, 1991a, 1991b), Boyd (1991), and Cooper (2004, 74–78).
- [12] My claim here rests on an oversimplified assumption regarding the ways in which looping effects operate. While my analysis suggests that looping effects express themselves as cultural *differences*, it is possible that some of the uniformities seen across cultures reflect particularly widespread *shared forms of interactivity* (e.g., global norms for expressing anxiety). Hence, looping effects can also manifest themselves as trans-cultural *similarities*. For the purposes of this paper, I would say that that the uniformity of a condition across cultures can *typically* be understood as a manifestation of (stable) biological aspects of that condition.
- [13] This issue complicated by the issue of ‘biolooping’ (Hacking 1999, 109–110). Biolooping is a phenomenon wherein changes in mental states result in changes in physiological states. For the purposes here, I assume that biolooping will not effect the *typical physiological patterns* associated with conditions such as schizophrenia or depression.

References

- APA (American Psychiatric Association). 2000. *Diagnostic and statistical manual of mental disorders*. 4th ed. (text rev.). Washington, DC: American Psychiatric Association.
- Bogen, James. 1988. Comments on ‘The sociology of knowledge about child abuse’. *Noûs* 22: 65–66.
- Boyd, Richard. 1991. Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies* 61: 127–148.
- Carlson, Neil R. 1999. *Foundations of physiological psychology*. 4th ed. Boston, MA: Allyn and Bacon.
- Cooper, Rachel. 2004. Why Hacking is wrong about human kinds. *British Journal for the Philosophy of Science* 55: 73–85.
- . 2005. *Classifying madness: A philosophical examination of the Diagnostic and statistical manual of mental disorders*. Dordrecht: Springer.

- Davison, Gerald C., and John M. Neale. 1996. *Abnormal psychology*. 6th ed. New York: John Wiley.
- Ereshefsky, Marc. 2004. Bridging the gap between human kinds and biological kinds. *Philosophy of Science* 71 (Proc.): 912–921.
- Foucault, Michel. 2006. *History of madness*. Translated by Jonathan Murphy and Jean Khalfa. London: Routledge. Original work published as *Folie et déraison: Histoire de la folie à l'âge classique* (Paris: Plon, 1961).
- Hacking, Ian. 1986. Making up people. In Hacking, *Historical ontology*. Cambridge, MA: Harvard University Press, 2002.
- . 1990. Natural kinds. In *Perspectives on Quine*, edited by Robert Barrett and Roger Gibson. Oxford: Blackwell.
- . 1991a. A tradition of natural kinds. *Philosophical Studies* 61: 109–126.
- . 1991b. On Boyd. *Philosophical Studies* 61: 149–154.
- . 1995a. The looping effects of human kinds. In *Causal cognition: A multidisciplinary debate*, edited by Dan Sperber, David Premack, and Ann J. Premack. Oxford: Clarendon Press.
- . 1995b. *Rewriting the soul: Multiple personality and the sciences of memory*. Princeton, NJ: Princeton University Press.
- . 1999. *The social construction of what?* Cambridge, MA: Harvard University Press.
- . 2004. Between Michel Foucault and Erving Goffman: Between discourse in the abstract and face-to-face interaction. *Economy and Society* 33: 277–302.
- . 2007. Kinds of people: Moving targets. *Proceedings of the British Academy* 151: 285–318.
- Haslam, Nick. 2002. Kinds of kinds: A conceptual taxonomy of psychiatric categories. *Philosophy, Psychiatry, and Psychology* 9: 203–217.
- Howard, George S. 1985. The role of values in the science of psychology. *American Psychologist* 40: 255–265.
- 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.
- Laing, R. D. 1967. *The politics of experience*. New York: Ballantine.
- McKim, William A. 2000. *Drugs and behavior: An introduction to behavioral pharmacology*. 4th ed. Upper Saddle River, NJ: Prentice Hall.
- Millon, Theodore, Paul H. Blaney, and Roger D. Davis, eds. 1999. *Oxford textbook of psychopathology*. New York: Oxford University Press.
- Murphy, Dominic. 2001. Hacking's reconciliation: Putting the biological and sociological together in the explanation of mental illness. *Philosophy of the Social Sciences* 31: 139–162.
- . 2006. *Psychiatry in the scientific image*. Cambridge, MA: MIT Press.
- Putnam, Hilary. 1975. The meaning of 'meaning'. In *Mind, language and reality: Philosophical papers*. Vol. 2. Cambridge: Cambridge University Press.
- Scheff, Thomas J. 1963. The role of the mentally ill and the dynamics of disorder: A research framework. *Sociometry* 26: 436–453.
- Schmaus, Warren. 1992. Sociology and Hacking's trousers. In *PSA 1992*. Vol. 1, edited by David Hull, Mickey Forbes, and Kathleen Okruhlik. East Lansing, MI: Philosophy of Science Association.
- Szasz, Thomas S. 1974 [1961]. *The myth of mental illness: Foundations of a theory of personal conduct*. Rev. ed. New York: Harper & Row.
- Thagard, Paul. 2003. Pathways to biomedical discovery. *Philosophy of Science* 70: 235–254.
- Zachar, Peter. 2000. Psychiatric disorders are not natural kinds. *Philosophy, Psychiatry, and Psychology* 7: 167–182.