Philosophy of Medicine
Original Research

What We Argue about when We Argue about Disease

Harriet Fagerberg[1]

1 London School of Economics and Political Science, London, United Kingdom. Email: h.fagerberg@lse.ac.uk

Abstract

The disease debate in philosophy of medicine has traditionally been billed as a debate over the correct conceptual analysis of the term "disease." This paper argues that although the debate's participants overwhelmingly claim to be in the business of conceptual analysis, they do not tend to argue as if this is the case. In particular, they often show a puzzling disregard for key parameters such as precise terminology, linguistic community, and actual usage. This prima facie strange feature of the debate points to an interesting and potentially instructive hypothesis: the disease debate makes little sense within the paradigm of conceptual analysis but makes good sense on the assumption that pathology is a real kind.

1. Introduction

What is disease? What does it mean for some condition to be pathological as opposed to normal and healthy? These are the sorts of questions addressed by the disease debate in the philosophy of medicine. Traditionally, it has been billed as a debate about the correct conceptual analysis of disease, with the aim of settling on a set of descriptive criteria, stated in terms of necessary and sufficient conditions, which are associated with the label "disease," and guide its usage in language. Definitions are proposed, counterexamples are concocted, alternative definitions proposed, and on it goes.

This paper argues that although participants in the disease debate overwhelmingly *claim* to be in the business of conceptual analysis, they do not tend to argue as if this is the case. In particular, they often show a puzzling disregard for key parameters—such as the terminology used, the linguistic community of interest, and the ways in which the relevant terms are, as a matter of fact, applied in the relevant contexts. The debate, in its current form, I argue, is confused. But the nature of the confusion points to an interesting and potentially instructive hypothesis. As I suggest, the style of argumentation employed by the debate participants makes little sense within the paradigm of conceptual analysis but makes good sense on the assumption that pathology is a real kind.

Before introducing the debate, some terminological preliminaries. First, within philosophy, "disease" has traditionally been understood *very* broadly to include all legitimately pathological conditions—from physical injury to disability to psychopathology. As some have pointed out, it is not entirely clear that this general usage exists outside of

ISSN 2692-3963 | Date Received 31-08-2023 | Date Accepted 14-11-2023 | Date First Published 21-12-2023
Philosophy of Medicine | DOI 10.5195/pom.2023.172 | Volume 4 | Issue 1 | pp.1–20

philosophy (Wakefield 2014; Cooper 2020). I comment on this in section 5. For now, I will use the term "disease" in this broad sense in accordance with its usage in the philosophical literature.

Second, I use "concept" in accordance with the assumption that conceptual analysis is—or could be—the analysis of concepts. In other words, I use "concept" to mean the mental states that guide the usage of a term. This does not mean that I am committed to this view of concepts; merely that I am deferring to the meta-semantics most friendly to conceptual analysis and the assumptions prevailing in the disease debate.

2. The Disease Debate in Philosophy of Medicine

Historically, the disease debate has focused on whether the correct conceptual analysis of disorder is evaluative (evaluativists) or, at least to some significant extent, rooted in objective biological facts (naturalists).[1] In what follows, I briefly outline the positions defended by four key participants.

We start with two well-known evaluativists. According to Lennard Nordenfelt: "A is completely healthy if and only if A is in a bodily and mental state which is such that A has a (second order) ability to realize all his or her vital goals, given accepted circumstances" (1995, 212). Health thus entails positive evaluation of the person as a whole (2007; 2018). "Vital goals" describe the goals whose attainment is necessary for the person's "minimal happiness"—meaning they are idiosyncratic to the individual. Illness, in turn, is the depletion of this ability. According to Nordenfelt, most cases of illness are caused by "malady"—a collective term referring to the full range of medical conditions. However, he leaves open the possibility that illness (that is, interference with one's ability to pursue one's vital goals) may be caused by things that are not in themselves "maladies" (2001). Nordenfelt describes his project as traditional conceptual analysis (1995).

According to Rachel Cooper, disease is: (1) a "bad thing to have;" (2) "that is such that we consider the afflicted person to have been unlucky;" and (3) "that can potentially be medically treated" (2002, 271). The first criterion is required in order to distinguish the pathological from the merely abnormal; the second to distinguish the pathological from the merely unpleasant; and the third to exclude nonmedical types of adversity, such as bankruptcy. She offers her analysis as a counter-thesis to Christopher Boorse's biostatistical theory, stating that "no biological account of disease can be provided because this class of conditions is by its nature anthropocentric and corresponds to no natural class of conditions in the world" (2002, 271). Cooper describes her evaluativist thesis as an analysis of what people have in mind in employing a term: "If a successful analysis was developed and presented to language users, they should recognize it as articulating their current practices (in the sense of 'Ah yes, that's what I was getting at!')" (2020, 143–144).

Attempting to bridge the naturalist–evaluativist divide, Jerome Wakefield influentially defends a mixed "harmful dysfunction" account (HDA). The concept of disorder, he contends, is jointly composed of: (1) "harm or deprivation of benefit to the person as judged by the standards of the person's culture;" and (2) the "inability of some internal mechanism to perform its natural function, wherein a natural function is an effect that is part of the

[1] Proponents of a value-laden analysis are usually collectively referred to as "normativists." There is, however, a sense in which the naturalists are also normativist—just in regard to biological, rather than social, norms. The label "evaluativist" is less confusing in this respect.

evolutionary explanation of the existence and structure of the mechanism" (1992, 384). These conditions are, on his view, individually necessary and jointly sufficient. Wakefield has always been explicit that the HDA is intended to be a "conceptual analysis" of "what we mean" when we say that some condition is a disorder (2007, 149). Conceptual analysis, he explains, should not be confused with "substantive scientific questions," such as discovering the underlying properties of a natural kind (2014, 652). Wakefield's original focus was on the concept of "disorder" as a constituent of "mental disorder" (1992). He also employs the supposedly coextensive term "medical disorder" (2014).

At the far naturalist end, we find Boorse's influential biostatistical theory of disease (BST). In his influential 1977 paper, Boorse defines health as "the readiness of each internal part to perform all its normal functions on typical occasions with at least typical efficiency" (1977, 555). On this basis, he defines disease as "a type of internal state which impairs health, i.e., reduces one or more functional abilities below typical efficiency" (2014, 684). Disease, he contends, is a "value-free scientific notion" (2014, 683). From the very beginning, Boorse described his project as an analysis of an existing concept (1975; 1977). In later work, he favors the term "pathological condition" over "disease," although his definition remains broadly unchanged.[2]

3. Methodology in the Disease Debate

Participants in the disease debate overwhelmingly claim to be in the business of "conceptual analysis." What does this mean exactly? Conceptual analysis is generally taken to be something like the search for psychological conditions or criteria that guide the application of a term among competent, fully informed members of some linguistic community. According to Karen Neander:

> Conceptual analysis is an attempt to describe certain features of the relationship between utterances of the term under analysis, and the beliefs, ideas, and perceptions of those who do the uttering. It involves trying to describe the criteria of application that the members of the linguistic community generally have (implicitly or explicitly) in mind when they use the term. (1991, 170)

Most often—in the disease debate and in general—a conceptual analysis is given in the form of necessary and sufficient conditions. As Maël Lemoine writes of the disease debate, the criteria "must be expressible under the form of necessary and sufficient conditions, on the one hand, and exceptions to these conditions, on the other" (2013, 310). As Neander argues, there is no inherent link between conceptual analysis and necessary and sufficient conditions (1991). The psychological literature tells us that rather than concepts being applied in accordance with neat criteria, people in fact use a variety of mental mechanisms, such as prototypes, exemplars, and similarity metrics, which are unlikely to be captured by itemized descriptions (Machery 2009). It is possible to study the mental states we associate with a term or concept without demanding that these states be expressible in the form of necessary and sufficient conditions—in the vein of Lemoine. However, participants in the

[2] In more recent work, Boorse entertains the possibility of making the BST relative to developmental stage, instead of chronological age (2014).

disease debate have overwhelmingly argued as if necessary and sufficient conditions are a prerequisite for successful analysis (see Lemoine 2013; Cooper 2020). In service of the main argument pursued in this paper, I will not seek to undermine this assumption.

How is conceptual analysis distinct from other philosophical methods? As Wakefield reminds us above, a conceptual analysis is *not* an attempt to arrive at substantial empirical truths about phenomena in the external world. In this sense, conceptual analysis contrasts with attempts at "real definition." A real definition seeks to describe underlying properties of kinds or phenomena—akin to a scientific theory (see Millikan 1989; 2017).[3] Where a real definition tells us, for example, what essential property underlies the natural kind gold (atomic number 79), conceptual analysis tells us something about what *we* have in mind in thinking about gold.

Moreover, conceptual analysis is a descriptive—rather than normative—matter. It is an attempt to say something about what members of the relevant linguistic community, as a matter of empirical fact, have in mind when they apply a term. It is not an attempt to say something about what they *should* have in mind, or how they *should* apply the term. Conceptual projects of such "normative" or "stipulative" varieties fall under the broad umbrella of what has recently come to be known as "conceptual engineering" or "conceptual ethics," and are considered an alternative to conceptual analysis, subject to distinct desiderata.

4. The Rules of Engagement

Given the characterization of conceptual analysis outlined above, what are the "rules" implicit in this methodology? What variables and restrictions should the conscientious, competent, and methodologically principled conceptual analyst be concerned with and pay attention to? If conceptual analysis is the search for necessary and sufficient conditions for the correct application of some term, there are three variables that should matter to the conceptual analyst: (1) the terms used; (2) the precise linguistic context of this usage; and (3) how the terms are, as a matter of fact, applied by competent, fully informed members of the relevant linguistic community. In what follows, I consider each in turn.

4.1 Words Matter

If conceptual analysis is the analysis of the criteria for the application of some existing term in a given linguistic community, it matters what exact word one takes oneself to be analyzing, and it matters that this word is sufficiently embedded in usage within the relevant community. One cannot simply assume that two or more terms are used synonymously or, at least, any claim to this effect would need careful substantiation. Moreover, if the term under analysis is a highly specialist philosophical term, or simply rarely used, it may not be

[3] Max Deutsch has recently argued that, properly understood, conceptual analysis also aims to describe external world phenomena. On this view, so-called conceptual analysis has nothing to do with concepts or representations per se, and is instead a confusing misnomer for something more akin to real definition (Deutsch 2021; see also Papineau 2009). For present purposes, I take traditional conceptual analysts at their word, and assume that conceptual analysis is the practice of describing the psychological criteria or mental representations that guide members of some linguistic community when applying some term or concept. I take it that this is both the most widely accepted view of conceptual analysis and also the view that accords best with what philosophers of medicine explicitly say about what they are up to.

sufficiently well integrated into ordinary, nonphilosophical usage to permit a well-evidenced conceptual analysis. For example, one might think there is little to learn about the concept of "emergence" from analyzing its usage among laypeople.

4.2 Context Matters

If conceptual analysis is the analysis of the application of a term in some linguistic context or community, the precise linguistic context matters. In other words, the principled conceptual analyst would carefully define the target linguistic community and the time period of interest, in the knowledge that even a successful conceptual analysis (one that *truly* reflects the criteria language-users have in mind when applying the term under analysis) is only successful relative to these parameters. In other linguistic communities at other times in history, the relevant term may have been used differently.

4.3 Usage Matters

Since conceptual analysis seeks to descriptively identify the conditions for the application of a concept in some linguistic context, the application of the concept in question is also "fixed" by its actual usage by competent, fully informed members of the target linguistic community.[4] This thus limits the scope for revisionism about how the concept *should* be used. If what you seek to explain is what *causes* the concept to be applied as it is, competent, informed application of the target concept in defiance of your analysis is evidence against its explanatory adequacy. The intellectually honest thing would then be for the originator to revise the proposed conceptual analysis—*not* the proper application of the target concept.

The question from here becomes: do participants in the disease debate generally heed these rules of engagement?

5. Methodological Oddities of the Current Debate

As argued above, the competent, methodologically principled conceptual analyst pays close attention to terminology, context, and application. In the following sections, I highlight some features of the disease debate in its current form that seem at odds with the rules defined above. In section 5.1., I consider participants' puzzling disregard for terminology. Next, I examine the inconsistent commitments of the debate's participants with regard to linguistic contexts. Lastly, I discuss their revisionism and tendency to be more attached to their proposed criteria than makes sense under conceptual analysis. In section 6, I make the case that these tendencies, though puzzling under conceptual analysis, make good sense if we assume that pathology is in fact a real kind.

5.1 The Irrelevance of Terminology

As argued, careful attention to terminology is a prerequisite for successful conceptual analysis. Consider how confused an analysis of "knowledge" would be if philosophers insisted that it must also account for related terms, such as "belief," "confidence," or

[4] This view of reference is what Millikan dismissively calls "conceptionism" (2000, 42).

"intuition." Such an analysis is not likely to be successful, precisely because these words are used in significantly distinct ways. Accordingly, careful attention to terminological differences is needed to ensure that we do not conflate distinct concepts.

Nevertheless, participants in the disease debate often seem strangely uninterested in terminological differences, abandoning or adopting terminology as it suits them, as if they could know—without further argument or evidence—that their analysis of some term A applies unaltered to a distinct term B. In what follows, I review the literature, with a focus on the views of the four philosophers introduced in section 2, and give examples of this tendency in action.

First, let us note the diversity of terms employed by conceptual analysts in the disease debate. For Nordenfelt, the target is the concept of "malady" or "ill health" (2018), while Cooper offers an analysis of, first, "disease" (2002) and, later, "disorder" (2020). Originally, Boorse described the BST as a conceptual analysis of "disease" (1977) but later pivoted to "pathological condition" (1997). Wakefield has focused on "mental disorder" or, rather, the disorder part of "mental disorder" (1992), or the broader term "medical disorder" (1997; 2014). All these philosophers maintain, however, that they are in the same business and talking about the same thing. Moreover, despite targeting different terms, participants in the disease debate tend to insist, without much additional argument, that their analyses apply unaltered to related words, often those words targeted by their competition (see, for example, Boorse 2011; Nordenfelt 2018; Wakefield 2014).

Why do participants in the disease debate think they are justified in generalizing their analyses across seemingly quite different terms, which could plausibly have quite different usage conditions, without substantial empirical justification? They might say that they are all analyzing a very general, broad notion of all things pathological or unhealthy—as I term it, the "general disease concept."

This tradition appears to originate in Boorse's initial defense of the BST. In his 1977 paper, Boorse proposed the BST as an analysis of "disease," understood "very broadly for any sort of condition violating perfect health, including injuries, poisonings, growth disorders, static defects, environmental stresses, and so on" (2011, 26). However, as critics have pointed out, it is not clear that this very broad concept (spanning everything from broken legs to personality disorders) is robustly embedded in any nonphilosophical usage. As Wakefield explains, Boorse was forced to give up on this general notion because "this usage is at best quite rare, and there is a question whether it exists at all other than as a convenient shorthand used in a few manuals and in the phrase 'health is the absence of disease'" (2014, 653). In response, Boorse amended his target from "disease" to "pathology," or "pathological condition," as used by medical pathologists (see Boorse 1997).

Perhaps Wakefield is right that Boorse is analyzing a concept that has no strong root in nonphilosophical usage. However, it seems an odd line of attack for Wakefield to pursue, as he would appear to be vulnerable to similar objections. Where does his very general concept of "medical disorder" originate? According to Wakefield, medical disorder is a "a term designed to capture the entire range of health deficits that are the essential target of medicine" (2014, 653). But it is not quite clear who uses the term "medical disorder" in this way. "Disorder" *simpliciter* appears mainly to be used in mental health (most prominently within the *Diagnostic and Statistical Manual of Mental Disorders,* or DSM) and in reference to noncontagious conditions (for example, endocrine disorders). The term "medical disorder" appears to be used relatively rarely—indeed, a search for "medical

disorder" on Google Scholar soon leads one to the work of a certain Jerome Wakefield. As such, if "disease" is not sufficiently integrated into ordinary usage to be worth analyzing, it is hard to see how Wakefield's concept of "medical disorder" could do the trick. Perhaps his vulnerability in this regard is *why* he describes "medical disorder" as having been "*designed*" (2014, 653; emphasis added). Perhaps this label is not as a matter of empirical fact applied in this way but is, instead, stipulated by design to cover the relevant extension. But designed by whom? This move—more conceptual engineering than conceptual analysis—is at odds with Wakefield's stated methodological aim.[5]

Nordenfelt, who takes himself to be analyzing the term "malady," is also committed to the existence of the general disease concept and, like Cooper, he positions his analysis in opposition to Boorse (Nordenfelt 1995; 2007; 2018). But where does this general concept come from? Nordenfelt does not appeal to laypeople's or medical professionals' usage but to that of philosophers: "'Malady' is a term that was introduced into the modern philosophical debate … by the philosophers Culver and Gert … in order to have an umbrella term for all the negative medical states" (2018, 15; see also Culver and Gert 1982). If "malady" is not embedded in nonphilosophical usage, it is a very questionable target for a conceptual analysis—unless, of course, we are analyzing its usage by *philosophers,* which seems a strangely circular project.

In a 2020 paper, Cooper makes what seems to me a remarkable concession (for a committed conceptual analyst): "Problematically, there is no common ordinary English word that covers the umbrella concept of 'any type of pathological state' and so different authors have ended up employing different terms" (2020, 143). If the term we are interested in does not really exist in nonphilosophical usage, this represents a significant challenge to the viability of traditional conceptual analysis in this area.

The point I wish to emphasize, for now, is not that conceptual analysis *is* unfeasible; rather, I want to draw attention to how cavalier philosophers have been in generalizing their analyses across (seemingly) distinct terminology, and how the uncertain status of the "general disease concept" has failed to deter those purporting to analyze its usage. In other words, although *words matter* for successful conceptual analysis, philosophers of medicine have tended to argue as if this were not the case.

5.2 Playing Fast and Loose with Linguistic Context

In failing to sufficiently factor in linguistic context, the conceptual analyst risks providing a single analysis for distinct concepts because patterns of usage differ. Therefore, the good conceptual analyst pays close attention to and carefully defines her particular linguistic community and time period of interest. Nevertheless, participants in the disease debate have often been vague, inconsistent, and ad hoc as to which linguistic community they have in mind.

[5] From what I can gather, Wakefield's usage of "medical disorder" in this way seems to follow from Robert Spitzer and Jean Endicott's 1978 attempt at defining mental disorder. They noted then an explicit commitment to viewing mental disorder as a subtype of the broader concept of "medical disorder" (Spitzer and Endicott 2018). This, in turn, may reflect a commitment to thinking of mental disorder in a way that is compatible with the medical model's view of psychiatric disorders as *real* medical problems—a counter to the anti-psychiatric contention that they are "myths" or mere problems in living (see Szasz 1960). This is arguably a normative constraint on what our concept of disorder *should* be, rather than a descriptive fact rooted in usage, which an adequate conceptual analysis must accommodate.

In a 2000 paper, Wakefield describes the HDA as "an account of disorder as it is used in medicine in general, including the mental health professions" (2000, 253). That is, he appears to specify the linguistic community of interest as professionals in the medical sphere (broadly construed)—doctors, nurses, psychiatrists, and so on. It appears that Wakefield is analyzing a clinical or scientific concept of disorder—certainly, a technical rather than a colloquial or lay concept. In light of this, it is a little perplexing that he is actively opposed to how the concept of disorder is in fact applied in medicine. Wakefield is extremely critical of what he takes to be "false positives" in disorder attribution, particularly within the DSM (Wakefield 1992; 1997; 2007; Horwitz and Wakefield 2007).

Challenging him on his apparent opposition to the prevailing scientific consensus, Dominic Murphy and Robert Woolfolk accuse him of covertly analyzing a "folk-concept" of disorder, and then defending it against whatever he intuits to be medical overreach. This, they contend, is not an attractive project: "Wakefield's position is the same as that of someone who claims to be analyzing the biologist's concept of mammal and then argues that our intuitive biological sense is that whales are fish, so the biological concept of mammal should not apply to them" (Murphy and Woolfolk 2000, 272). So, is Wakefield analyzing the concept as it is used by laypeople, or as it is used by medical professionals?

By 2014, Wakefield is quite explicit that he includes both lay and scientific linguistic communities in his analysis of "medical disorder," stating that "clinicians … form the bulk of the medical community and *prima facie* are the most plausible target for a conceptual analysis of 'medical disorder'" (2014, 651) but, also, in an analysis "of the concept of disorder, laypeople should be included in the linguistic community from which evidence is drawn" (2014, 652). However, by 2021, he remarks: "Precisely specifying a preset target community is not necessary" because "the 'community' is in effect a construct of those dispersed throughout the professional and lay communities who share a certain widespread concept of disorder that is a salient one among the many in circulation and has certain properties that make it important in scientific research and lay debate" (2021, 139). Wakefield's position suddenly seems strangely circular. "The community" are those who possess the concept, as he understands it?

For now, note that Wakefield's methodological commitments, as pertaining to the linguistic community of interest, seem surprisingly ad hoc. If you are really interested in mapping the criteria for the application of some particular term within some particular linguistic context, it makes *all* the difference which linguistic community you are targeting. However, Wakefield writes as if these are details to be worked out as and when—or, indeed, that may not matter at all.

Unlike Wakefield, Boorse has always maintained that he is analyzing a technical, scientific concept reflected in specialist medical usage. In his early writing, he described the BST as an analysis of the concept of "disease" as it appears in the American Medical Association's *Standard Nomenclature of Diseases and Operations,* "the discursive context of medical textbooks and research papers," and in the phrase "health is the absence of disease" (1977, 552). He explicitly excludes laypeople from his linguistic community of interest: "[The BST] aims to analyze a theoretical medical concept of disease in the sense of 'pathological condition' … and for this task lay ideas of disease are of little interest" (2014, 688). He also excludes usage in psychiatry and mental health fields (1977).

Following his initial defense, Boorse was soon confronted with a seemingly troubling implication. If disease is simply subnormal part-function, it would seem we are all diseased

and none of us healthy, for there are no doubt one or more dead cells within our otherwise healthy bodies, which thus compromise our perfect Boorsian health (see Nordenfelt 1987). In response, Boorse replies: "My basic analysis really aims at the pathologist's, not the clinician's, concept of disease" (1997, 17) and "I am content for the BST to live or die by the considered usage of pathologists" (1997, 53). Furthermore, "to the pathologist no one is normal" (1997, 50). This is a significant departure for Boorse, as the notion of "universal disease" is arguably *not* the notion of "health is the absence of disease"—the usage that initially motivated his inquiry.[6] Nor will you find "single dead cell within an otherwise healthy body" listed within medical classifications. In other words, when his analysis fails to chime with the usage that supposedly underpins it, Boorse just adopts a new target community.

Nordenfelt, in contrast to Boorse, claims to be analyzing an "ordinary lay notion of health" (2018, 12). However, this does not deter him from positing his theory in explicit opposition to Boorse's BST: "I will criticize Boorse's proposal mainly by using arguments from the analysis of ordinary lay and medical language" (2018, 10). Given that Boorse has a different community of language-users in mind—medical pathologists—Nordenfelt's objections do not make much methodological sense.[7] Boorse would be the first to concede, as we have seen, that his theory may fail to make sense of ordinary lay or clinical notions. If Nordenfelt were arguing in accordance with the rules, he would recognize that conceptual analysis is relative to context, and thus may simply vary in accordance with variance in this parameter without prejudice to either theory. Objecting to the BST on the grounds that it does not accord with lay usage is faulting an apple for being a bad orange.

The unsubstantiated imposition of an analysis intended to account for a concept's application in linguistic community A onto some other linguistic community B does not make much sense from the perspective of conceptual analysis. If your interest is truly in the criteria for the application of some term within some community, it makes sense to start with the community—not the criteria. In sum, participants in the disease debate often show a surprising lack of interest in manner in which their analyses are relative to linguistic context.

5.3 Usage, Revisionism, and "Being Too Attached"

In doing conceptual analysis, usage—how the term is *in fact* applied by competent, fully informed members of the relevant linguistic community—matters. If your interest is in the necessary and sufficient criteria for the application of the term, the term's extension (that is, the set of cases an analysis must aim to capture) is also "fixed" by its competent, informed application. The ability of a conceptual analysis to successfully fence in the cases to which the term is applied is thus the primary measure of its success (see also Lemoine 2013).

Because of this, the disease debate has largely been a game of counterexamples. For example, one might object to Cooper's analysis of disease on the grounds that an unwanted pregnancy—being bad, unlucky, and potentially medically treatable—counts as a disease on her view. Surely, pregnancy (however undesirable) is a paradigm case of healthy, normal bodily operation? Faced with such a counterexample, the embattled conceptual analyst

[6] See also Wakefield on this point: "Boorse does not show that the pathologist would judge that pathology to constitute the individual's deviation from ideal health" (2014, 657).
[7] Nordenfelt is well aware that Boorse is targeting a technical notion of disease (see Nordenfelt 2018; 2001).

might be tempted to retort that even though the term "disease" is not generally applied to pregnancy, properly understood, unwanted pregnancy falls within the extension of the term of interest—precisely because it satisfies her analysis:

> I suggest that our intuitions do not cohere with my account here because our intuitions as to whether or not a condition is a disease lag behind changes in the disease-status of a condition … Prior to the invention of effective contraceptives those who had pregnancies they would have been better off without were not unlucky, and until comparatively recently it has been socially unacceptable to treat unwanted pregnancy … We still think of unwanted pregnancy as not being a disorder because our intuitions lag behind changes in the disorder-status of a condition. (Cooper 2002, 278–279)

Cooper suggest that we are undergoing a change in the term's application that will eventually vindicate her analysis. She offers little evidence for this, beyond the fact that unwanted pregnancy (now) satisfies her analysis of disease, while before it did not. What reason do we have to think that such a change is underway, beyond its convenient accord with Cooper's argumentative objectives? (We might be tempted to note, twenty years later, that no such conceptual change appears to have occurred.)

For Boorse, the inclusion of pregnancy in the category of disease is quite obviously the wrong result: "It is the analytic equivalent of the 'Game Over' sign" (1997, 44). However, Boorse himself is not above stipulation. His original formulation of the BST contained a clause intended to account for common—or near-universal—diseases, such as tooth decay, which are not abnormal, and thus would not otherwise count as pathological on the BST. However, he later abandons the clause, on the grounds that it is not "worth the criticism it evoked" (2014, 684). In other words, Boorse would rather just stipulate that tooth decay is—contrary to usage—non-pathological because it does not satisfy the BST.

Similarly, Boorse has insisted that the functional degradation accompanying old age (and resulting, invariably, in the organism's death) is not pathological: "When senile decline of function is caused from within, our account will not allow it to be a disease" (1977, 567). He affirms this contention in 2002, concluding that it is not the BST that is faulty—rather, it is the judgments of medicine itself. Although medical specialists would likely consider atherosclerosis and prostate cancer to be pathological states, regardless of the age group afflicted, in this "medicine is wrong … what is pathological is only age excessive atherosclerosis, premature prostate cancer, and so on" (2002, 103). Again, Boorse seems oddly comfortable dictating the usage of a term he claims merely to be analyzing.[8]

Despite the incompatibility of revisionism with conceptual analysis, participants in the disease debate often seem keen to *rule on* (rather than merely capture) the extension of the concept of interest. Notably, over the last few decades, Wakefield has argued forcefully against "false positives" in psychiatric diagnosis. His concern, in short, is that DSM diagnoses increasingly stretch beyond the legitimate realm of medicine and into the realm of "normal" healthy functioning. More precisely, Wakefield worries that an increasing number of the instances to which the label "disorder" is as a matter of fact applied by the community fail to constitute evolutionary dysfunctions (and thus fail to satisfy the HDA).

[8] It should be noted that he has rowed back on this contention in more recent writings (see Boorse 2014).

For example, Wakefield contends that psychiatry has hijacked and appropriated "normal sorrow" and "natural anxieties," which may be adaptive (Horwitz and Wakefield 2012; 2007). He also has concerns as to the "conceptual validity," as he terms it, of specific disorder typologies. He expresses doubt as to whether "binge eating disorder" entails dysfunction: "It remains unclear why the common tendency to overeat when food is plentifully available … is classifiable as a disorder versus a normal variation of an evolutionarily shaped inclination to amass scarce calories" (2016, 122). Similarly, he casts doubt on the status of social anxiety disorder (2007, 154–155), and expresses concern with the high rates at which depression and attention deficit hyperactivity disorder (ADHD) are diagnosed (2015, 190).

What is this "medical reality" Wakefield defends against psychiatric overreach? The term he claims to be analyzing—disorder—is most frequently applied in psychiatry and its classifications but he cannot be leaning here on a technical psychiatric conception of "disorder," as psychiatry is, after all, the culprit. If Wakefield were analyzing a technical concept of disorder rooted in specialist psychiatric usage, he would have to fess up to the fact that in these specialist domains the term "disorder" is consistently applied in a way which (at his own insistence) eludes his proposed analysis.

My guess is that Wakefield wishes for the philosophical resources to say something more substantial than conceptual analysis can afford him. Even assuming that he is correct—and that laypeople and experts really do have harm and failure of a naturally selected effect in mind when applying the term "disorder"—this does not ground any stronger normativity than that of policing correct language use. Specialists, such as those within DSM working groups, can simply choose to use the term in a nonstandard way that fails to accord with normal judgments. And why should they not? Wakefield would need something more normatively meaty than a conceptual analysis to make this case.[9]

While already critical of DSM-IV, Wakefield seems particularly unhappy with a number of changes introduced in the DSM-5, which he sees as further liberalizing the conditions under which the disorder label can be applied. For him, this represents a threat to the very concept of disorder. This leads me to a final, puzzling characteristic of the disease debate. For lack of a better term, I call it "being too attached."

Suppose that Wakefield was, originally, analyzing the term "mental disorder," roughly as it features in specialist usage and classification, prior to the publication of DSM-5. In other words, he believes that dysfunction and harm are individually necessary and jointly sufficient for the application of the term "mental disorder" in psychiatry and mental health fields in the 1990s and early to mid-2000s. Suppose he is right about this. Then, in 2013, the American Psychiatric Association publishes the DSM-5. Upon inspecting its contents, Wakefield realizes that the term "mental disorder" is now applied rather more liberally. Specifically, he notes that the dysfunction criterion no longer appears to be necessary for

[9] Wakefield's view is that function talk refers to a particular evolutionary notion of biological function on which functions are phylogenetically selected effects (Wakefield 1999); perhaps this is what grounds and justifies his revisionism? I admit I find this feature of Wakefield's view somewhat inelegant. It has the effect of making his view hybrid in two ways—in combining facts and values, and in combining two different approaches to concepts. While the concept of function is a "black box essentialist concept," the conceptual content of which is simply "the things which share the underlying essence of psychological prototypes of function," the concept of medical disorder is determined instead by a neat list of descriptive criteria—that is, the HDA (for discussion, see Fagerberg 2023). More pertinently in the present context, this means that Wakefield's HDA comes with much of the methodological baggage, yet few of the benefits, of a theory of pathology as a real kind. At this point, why not just take the plunge?

the application of the term "disorder" among members of DSM working groups and, consequently, within the classification system itself. Many of the disorders included in DSM-5 are *not* evolutionary dysfunctions. How should Wakefield respond?

Of course, the conditions under which a term is applied are relative not only to the particular linguistic community of interest but also to the particular *time* of analysis. These conditions can simply *change*. As language, culture, and science progress, dysfunction can just cease to be necessary. Nothing about conceptual analysis guarantees stability. However, this is no threat to the principled, committed conceptual analyst who heeds the rules of engagement. If she is truly just interested in mapping the criteria for the application of the term "mental disorder" in some context, upon noting a change, the principled conceptual analyst just incorporates this new data, disinterestedly, into her analysis. Perhaps Wakefield was right, before, that dysfunction *was* necessary but times have since changed. The methodologically sound thing to do, when faced with this fact, would be to note a slight change and then move on. But instead, Wakefield seeks to "correct" this usage. He is, in my terminology, "too attached." He has married his preferred analysis in a way that transcends the logic of his methodology.

Similarly, in a recent paper, Cooper discusses an apparent conceptual change from DSM-IV to DSM-5. It seems, she argues, that "harm" has ceased to be necessary for the application of the term "disorder":

> Until very recently, I defended a traditional value-laden approach to disorder. I thought that a condition could only be a disorder if it is harmful for the individual with it. However … our concept of disorder is shifting. I think that it was necessary for a condition to be a disorder that it be harmful—but conceptual change means that this requirement has recently become questionable and may sometimes now be abandoned (by at least some linguistic sub-communities). (Cooper 2020, 152)

She rightly notes: "When a conceptual analysis finds a condition to be necessary for a term to be applied, this necessity may not be projectible into the future. The necessity depends on the current consensus of a linguistic community, and they might change their mind" (2020, 149).

Cooper is therefore worried that her argument for a value-laden conceptual analysis of disorder is collapsing. She rejects several solutions to the problem as it stands (giving up, revisionism) before embracing a third strategy, which she terms "belt and brace." In short, it entails "doubling up" her evaluativist claim. Cooper suggests that values may be relevant to disorder in more ways than one. Even if the necessity of evaluative harm is under threat from conceptual change, we could "belt and brace" the evaluativist claim by arguing that, nonetheless, values enter the picture in other respects. For example, we could *also* make the case, preemptively, that values enter into deciding when, along a continuum, normal function descends into dysfunction. Now that Cooper makes *two* distinct claims of value dependence, should one fail, she can plausibly rely on the other to support her overall evaluativist thesis:

> Those who want their trousers to stay up, but who have reason to doubt the reliability of their belt, and also have some suspicion about the strength of their braces, can maximize their chances of staying dressed by using both. Analogously, it may sometimes

> be possible to "'belt and brace" conceptually-grounded claims, by grounding a claim in multiple semi-independent considerations. (Cooper 2020, 152)

Cooper is not motivated here by getting the criteria of application in the relevant community exactly right—she herself acknowledges a change. Rather, she is motivated to rescue her overall evaluativist project by tinkering with her original claim. However, it is entirely unclear what justifies her attachment to this particular analysis. The principled conceptual analyst, heeding the rules of engagement, would just note that there has been a slight change in the criteria guiding the term's application, and then leave it at that—concept successfully analyzed. If your aim is just to map the conditions under which some term is as a matter of descriptive fact applied, conceptual change is no threat to your success; you can simply factor in the change and then move along. To continue Cooper's analogy, why not just drop your trousers?

Cooper's instinct to "belt and brace" is, I suggest, indicative of a general tendency among participants in this debate, exemplified here by Wakefield and Cooper, toward wanting to defend a stronger, more metaphysically substantial thesis than conceptual analysis can support. It makes little sense to have such an inflexible commitment to something so contingent as the conditions under which some specific community would apply a term. These are open to change, and why should this concern us?

6. Accounting for the Confusion

We have seen participants in the disease debate claim to be in the business of conceptual analysis but show a strange disregard for the rules implicit in their methodology. They are often imprecise or inconsistent as to which term or community they take as their target. Moreover, despite a lack of normative force in conceptual analysis, they often seek to stipulate application, and have been oddly insensitive to the fact that even a successful conceptual analysis is not impermeable to change. The disease debate, in its current form, seems strangely distant from its method.

Is the disease debate irredeemably confused? Perhaps. However, I wish to offer a hypothesis that makes sense of the nature of the confusion. The methodological oddities outlined above make *more* sense on the assumption that pathology is a real kind—a fact that tacitly influences the debate's participants and thus the character of the debate itself. In what follows, I first briefly explain what I take real kinds to be, before showing that these prima facie puzzling features of the disease debate can be made more sense of on the assumption that pathology is a real kind.

6.1 Real Kinds

According to an influential paradigm in philosophy, exemplified by the work of Ruth Millikan (2000; 2017), Richard Boyd (1999), Muhammad Ali Khalidi (2013; 2018), and others, real kinds are categories of instances that share a great many properties in common. In other words, real kinds are "property clusters." For example, instances of the kind *Panthera leo* share many features in common—they have claws, yellowish fur, a characteristic tufted tail, and so on. When properties co-occur across instances, it permits us to make generalizations which, in turn, support inferences. For example, if we note that

some particular large animal has a characteristic tufted tail, we may infer—not infallibly but with some reliability—that it is a lion and, accordingly, is probably carnivorous too.

According to Millikan, the fact that our world is composed of densely interlocked clusters of properties—as she puts it, the fact that our world is "clumpy"—is what supports everyday knowledge and induction (Millikan 2017). However, we do not need be *aware* that our world is composed of real kinds to make use of their rich inductive potential. Our brains are adapted to a "clumpy world" even if we have no explicit awareness of this being the case. In a similar vein, then, you might imagine that our philosophical debates—such as the disease debate—may be shaped by the clumpy structure of our world even if the debate's participants lack awareness of, or even deny, this structure. This is my claim: the character of the disease debate is shaped by the existence of a real kind in the world.

Those who take a cluster view of kinds often deny that real kinds are necessarily characterized by "essences."[10] As such, it is not a given that there will be anything essential in the world that would allow us to arrive at a precise characterization of the relevant kind. However, as Millikan acknowledges, philosophers do not typically count as real kinds groups of individuals that share properties for *no good reason*. That is, there is usually some non-accidental reason or causal structure in the world that explains why members of a kind share properties (see also Khalidi 2013; 2018). For example, samples of the chemical kind gold share lots of properties in common across instances—being malleable, being dense, being heat conductive—*because* they share the same atomic structure (Godman, Mallozzi, and Papineau 2020).

Against this backdrop, I want to introduce some further nuance. Millikan recognizes not just real kinds but also "real categories." Real categories are analogues of real kinds. Where real kinds are correlations of *determinate* properties in nature, real categories are correlations of determinables in nature or, more simply put, kinds of kinds (Millikan 2017). For example, "chemical element" is a real category composed of chemical kinds.[11] To be precise, my view is that pathology is a Millikanian "impure real category"—that is, a real category that is *also* a real kind. This metaphysical framework facilitates a unified account of pathology as a real kind, which in turn encompasses a number of more specific kinds (tuberculosis, scurvy, and so on).[12] Unpacking and defending this view in full goes beyond the scope of this paper. For now, I am simply concerned to motivate the position that the apparent confusion of the disease debate makes *more* sense on the assumption that pathology is (like) a real kind.

[10] As Millikan puts it: "Clusters … may fail to exhibit universal distinguishing properties so as to fall under necessary descriptions" (2017, 13).
[11] Notably, Millikan cites disease as an example of a "real category" (2017).
[12] In *The Nature of Disease* Lawrie Reznek argues forcefully that diseases or pathological conditions do not, as a class, constitute a natural kind (Reznek 1987). He leaves it open, however, that individual disease types form natural kinds. My view is that disease is a real or natural kind, individual disease kinds are natural kinds, and these two facts are not unrelated. Reznek correctly observes that it is hard to prove a negative; that is, to conclusively establish that pathological conditions fail to form a natural kind, through showing that no theory of the "essence" of pathology holds up (1995). Similarly, it is hard to prove that disease *does* form a natural kind without providing any good theory as to the casual structure of this kind. However, it is beyond the scope of this paper to provide such an account. For now, I am just concerned to provide initial motivation for such a view.

6.2 Understanding the Disease Debate

In the following sections I argue that each of the methodologically puzzling features of the disease debate outlined in section 5 can be understood better on the assumption that there is a real kind out there to which disease-talk refers, and which participants in the disease debate are tacitly influenced by.

6.2.1 Words, Context, and Conceptions

As we have seen, there is no real agreement in the disease debate as to which word is the target of analysis. However, this does not trouble the debate's participants—they all seem to know what it is they are arguing about, even if there is no definite word for it.

Nor is there any real consensus on which usage—or which community of language-users—the debate's participants are interested in. As we have seen, Wakefield includes both laypeople and medical professionals, while Boorse has in mind official classifications and medical pathologists. Cooper has lately focused her scholarship mainly on DSM working groups (Cooper 2018; Cooper and Blashfield 2016), while Nordenfelt's analysis is firmly rooted in ordinary lay usage (2018). Why assume that all these different linguistic subgroups are really talking about the same thing?

Neither is there much evidence of a consensus in people's representations of the subject matter. We have little reason to think there is a uniform general conception of disease that guides our usage of the family of medical terms which, supposedly, co-refer to cover the extent of the general medical terrain the disease debate concerns. Indeed, the very existence of the debate is evidence that this is not the case. Seemingly, people have very different things "in mind" when talking about disease.

In sum, there is no single word, pattern of usage, or set of beliefs that unifies our concept of the pathological. This raises the question: why are participants in the disease debate so sure that there is a *single* concept here to analyze? Why do they have such confidence that they are talking and arguing about the very same thing?

My suggestion is that while there may be no single word, no single pattern of usage, and no single conception that unifies our concept of the pathological, there *is* a uniform real kind "pathology" in the world. The debate's participants are so sure, despite no consensus on terminology, that they are talking and arguing about the very same thing because there is a uniform cluster of biological conditions out there in the world that share objective properties in common. It is the *referent* of our disease-talk that is uniform, not the language we use to refer to it. The contexts in which we can confidently say that distinct words express the very same concept are the contexts in which we use different words to refer to some uniform worldly referent, such as a real kind (see Sawyer 2020).

Similarly, we can expect the criteria for the application of a term to vary among linguistic communities. According to the logic of conceptual analysis, it makes all the difference which community we have in mind. However, we *do not* expect the defining properties of a real kind to vary among different linguistic subcommunities. If there is a uniform, real biological kind out there in the world, to which talk of the pathological (disease, disorder, and so on) refers, these disparate linguistic communities—whether psychiatrists at the Maudsley or medical pathologists in New York—are in a very real sense talking about the *same kind of thing*. The hypothesis that pathology is a real kind makes good sense of the debate's participants' perplexing disregard for precise linguistic context, and their insistence that

they are, in fact, arguing about the same thing—despite apparent disunity on all relevant metrics. Covertly, they are leaning on a notion of pathology that is unified not by our terminology, usage, or conceptions, but by how the world actually is.

6.2.2 Revisionism and "Being Too Attached"

As argued here, participants in the disease debate often display a puzzling revisionist bent, which lacks justification in conceptual analysis. If the aim of analysis is to capture how a word is applied, application in defiance of the proposed analysis is evidence against it. The conceptual analyst does not then have the philosophical resources to *order* the relevant linguistic communities to change their minds on the matter.

Relatedly, participants in the disease debate often seem wedded to their proposed analyses in ways that cannot be justified by their stated aims. If your aim is simply to understand what conditions guide the application of a term, a change in these conditions over time is no threat to your aim. It simply does not make sense, as Wakefield does, to oppose such linguistic change or, like Cooper, to "belt and brace" your conceptual claim to guard against its expiration. This stubborn attachment to particular views of what people "have in mind" when applying disease-terms cannot really be motivated by a desire to get the criteria of application right, so it is unclear what *does* motivate it. How do we make sense of the overattachment and revisionism of philosophers of medicine in this regard?

Let us start with revisionism. If your aim is to explain what guides the application of a term and your analysis fails to make sense of how people *in fact* apply the term, your analysis is bad at explaining what it seeks to explain. Thus, an application of, say, "disorder" that falls outside your analysis is a counterexample. However, suppose instead that pathology is a real kind. The boundaries of a real kind need not correspond exactly to how the relevant kind-term is applied in language (see Millikan 2000). Indeed, the linguistic community may be systematically mistaken as to whether some particular instance is an instance of the relevant real kind. Applying this to the disease debate, then, it might be that mental health professionals in the United States are systematically *wrong* about what sorts of mental conditions count as instances of pathology. Thus, if pathology is a real kind, we can make good sense of principled revisionism as regards the concept's proper application.

Now consider philosophers' overattachment to their proposed conceptual analyses. As noted, it does not make sense to mount a stubborn defense of your thesis as to the criteria for the application of some term when faced with good evidence that these conditions are *changing*. There is nothing about cultural, social, and linguistic practices that guarantees invariance over time. However, the properties associated with a real kind can be expected to be fairly stable.[13] It would make sense, then, to remain loyal to a particular theory of how the world actually is—that is, which causal structure underpins instances of pathology in the world—and to defend this theory against changing conceptions (ideas, beliefs, and so on) that fail to reflect your theory. My suggestion is that it is the presence of a real kind that motivates participants to engage in normative stipulation and to defend their accounts against change.

[13] There are exceptions to this rule. For example, some "interactive kinds" are real kinds but may yet change (relatively) quickly (see Cooper 2004; Khalidi 2013; Fagerberg 2022).

7. Conclusion

I have argued that a critical eye on the practices of philosophers participating in the traditional disease debate motivates the position that pathology is (like) a real kind. In turn, this also suggests a new approach to the disease debate itself. Exploring the hypothesis that pathology is a real kind, we may now sensibly ask a number of further questions. What are the properties that instances of pathology share in common? What is the underlying causal structure that accounts for the clustering of these properties in nature? As such, this hypothesis opens up a fruitful new avenue for further inquiry. On these grounds, I suggest we adopt this alternative working hypothesis and see what we can discover.

This hypothesis, if correct, does not invalidate other approaches to defining disease. It may be that a conceptual analysis of pathology, once properly specified, can be supplied. Or perhaps we can discover ways of thinking about disease that serve legitimate pragmatic and ethical aims in specific contexts. But we must be clear-eyed about the limitations of our methods, and take care not to conflate our diverse methodological aims and interests. If it is true, as I have suggested, that in some implicit sense philosophers of medicine have been confused as to whether they are providing a description of a real kind, or an analysis of what people "have in mind" when using disease-terms, the current debate holds little promise—either for a conceptual analysis or a real definition of disease. Clearly distinguishing our aims is an essential prerequisite for coherent debate and, ultimately, progress.

Acknowledgments

I would like to thank David Papineau, Alexander Bird, and Justin Garson for helpful feedback on earlier drafts of this paper. I would also like to thank members of the Work in Progress Seminars at King's College London and the Berlin School of Mind and Brain for their insightful thoughts and questions. Finally, I am grateful to the two anonymous reviewers for their input.

Disclosure Statement

No competing interest was reported by the author.

References

Boorse, Christopher. 1975. "On the Distinction between Disease and Illness." *Philosophy & Public Affairs* 5, no. 1: 49–68. http://www.jstor.org/stable/2265020.

———. 1977. "Health as a Theoretical Concept." *Philosophy of Science* 44, no. 4: 542–573. https://doi.org/10.1086/288768.

———. 1997. "A Rebuttal on Health." In *What Is Disease?,* edited by James M. Humber and Robert F. Almeder, 1–134. Totowa, NJ: Humana Press. https://link.springer.com/chapter/10.1007/978-1-59259-451-1_1.

———. 2002. "A Rebuttal on Functions." In *Functions: New Essays in the Philosophy of Psychology and Biology,* edited by Andre Ariew, Robert C. Cummins, and Mark Perlman, 63–112. Oxford: Oxford University Press.

———. 2011. "Concepts of Health and Disease." In *Philosophy of Medicine,* edited by Fred Gifford, 13–64. Oxford: North-Holland.

———. 2014. “A Second Rebuttal on Health.” *Journal of Medicine and Philosophy* 39, no. 6: 683–724. https://doi.org/10.1093/jmp/jhu035.

Boyd, Richard. 1999. “Homeostasis, Species, and Higher Taxa.” In *Species: New Interdisciplinary Essays,* edited by Robert A. Wilson, 141–185. Cambridge, MA: MIT Press.

Cooper, Rachel. 2002. “Disease.” *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 33, no. 2: 263–282. https://doi.org/10.1016/S0039-3681(02)00018-3.

———. 2004. “Why Hacking Is Wrong about Human Kinds.” *British Journal for the Philosophy of Science* 55, no. 1: 73–85. https://doi.org/10.1093/bjps/55.1.73.

———. 2018. “Understanding the DSM-5: Stasis and Change.” *History of Psychiatry* 29, no. 1: 49–65. https://doi.org/10.1177/0957154X17741783.

———. 2020. “The Concept of Disorder Revisited: Robustly Value-Laden despite Change.” *Aristotelian Society Supplementary Volume* 94, no. 1: 141–161. https://doi.org/10.1093/arisup/akaa010.

Cooper, Rachel and Roger K. Blashfield. 2016. “Re-evaluating DSM-I.” *Psychological Medicine* 46, no. 3: 449–456. https://doi.org/10.1017/s0033291715002093.

Culver, Charles M. and Bernard Gert. 1982. *Philosophy in Medicine: Conceptual and Ethical Issues in Medicine and Psychiatry*. Oxford: Oxford University Press.

Deutsch, Max. 2021. “Conceptual Analysis without Concepts.” *Synthese* 198, no. 11: 11125–11157. https://doi.org/10.1007/s11229-020-02775-0.

Fagerberg, Harriet. 2022. “Reactive Natural Kinds and Varieties of Dependence.” *European Journal for Philosophy of Science* 12, no. 4, article 72. https://doi.org/10.1007/s13194-022-00500-x.

———. 2023. “Medical Disorder Is Not a Black Box Essentialist Concept: Review of *Defining Mental Disorder: Jerome Wakefield and His Critics,* edited by Luc Faucher and Denis Forest.” *Philosophy of Medicine* 4, no. 1: 1–9. https://doi.org/10.5195/pom.2023.165.

Godman, Marion, Antonella Mallozzi, and David Papineau. 2020. “Essential Properties are Super-Explanatory: Taming Metaphysical Modality.” *Journal of the American Philosophical Association* 6, no. 3: 316–334. http://dx.doi.org/10.1017/apa.2019.48.

Horwitz, Allan V. and Jerome C. Wakefield. 2007. *The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder*. Oxford: Oxford University Press.

———. 2012. *All We Have to Fear: Psychiatry’s Transformation of Natural Anxieties into Mental Disorders*. Oxford: Oxford University Press.

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

———. 2018. “Natural Kinds as Nodes in Causal Networks.” *Synthese* 195: 1379–1396. https://doi.org/10.1007/s11229-015-0841-y.

Lemoine, Maël. 2013. “Defining Disease beyond Conceptual Analysis: An Analysis of Conceptual Analysis in Philosophy of Medicine.” *Theoretical Medicine and Bioethics* 34, no. 4: 309–325. https://doi.org/10.1007/s11017-013-9261-5.

Machery, Edouard. 2009. *Doing without Concepts*. Oxford: Oxford University Press.

Millikan, Ruth G. 1989. "In Defense of Proper Functions." *Philosophy of Science* 56, no. 2: 288–302. https://doi.org/10.1086/289488.

———. 2000. *On Clear and Confused Ideas: An Essay about Substance Concepts*. Cambridge: Cambridge University Press.

———. 2017. *Beyond Concepts: Unicepts, Language, and Natural Information*. Oxford: Oxford University Press.

Murphy, Dominic and Robert L. Woolfolk. 2000. "Conceptual Analysis versus Scientific Understanding: An Assessment of Wakefield's Folk Psychiatry." *Philosophy, Psychiatry & Psychology* 7, no. 4: 271–293. https://muse.jhu.edu/article/28365.

Neander, Karen. 1991. "Functions as Selected Effects: The Conceptual Analyst's Defense." *Philosophy of Science* 58, no. 2: 168–184. https://www.jstor.org/stable/187457.

Nordenfelt, Lennard. 1987. *On the Nature of Health: An Action-Theoretic Approach*. Dordrecht: D. Reidel.

———. 1995. *On the Nature of Health: An Action-Theoretic Approach*. Dordrecht: Springer Science & Business Media.

———. 2001. *Health, Science, and Ordinary Language*. Amsterdam: Rodopi.

———. 2007. "The Concepts of Health and Illness Revisited." *Medicine, Health Care and Philosophy* 10, no. 1: 5–10. https://doi.org/10.1007/s11019-006-9017-3.

———. 2018. "Functions and Health: Towards a Praxis-Oriented Concept of Health." *Biological Theory* 13, no. 1: 10–16. https://doi.org/10.1007/s13752-017-0270-x.

Papineau, David. 2009. "The Poverty of Analysis." *Aristotelian Society Supplementary Volume* 83, no. 1: 1–30. http://dx.doi.org/10.1111/j.1467-8349.2009.00170.x.

Reznek, Lawrie. 1987. *The Nature of Disease*. New York: Routledge.

———. 1995. "Dis-ease about Kinds: Reply to D'amico." *Journal of Medicine and Philosophy* 20, no. 5: 571–584. https://doi.org/10.1093/jmp/20.5.571.

Sawyer, Sarah. 2020. "Truth and Objectivity in Conceptual Engineering." *Inquiry* 63, no. 9–10: 1001–1022. https://doi.org/10.1080/0020174X.2020.1805708.

Spitzer, Robert L and Jean Endicott. 2018. "Medical and Mental Disorder: Proposed Definition and Criteria." *Annales Médico-psychologiques, revue psychiatrique* 176, no. 7: 656–665. https://doi.org/10.1016/j.amp.2018.07.004.

Szasz, Thomas S. 1960. "The Myth of Mental Illness." *American Psychologist* 15, no. 2: 113–118. https://doi.org/10.1037/h0046535.

Wakefield, Jerome C. 1992. "The Concept of Mental Disorder: On the Boundary between Biological Facts and Social Values." *American Psychologist* 47, no. 3: 373–388. https://doi.org/10.1037//0003-066x.47.3.373.

———. 1997. "Diagnosing DSM-IV—Part I: DSM-IV and the Concept of Disorder." *Behaviour Research and Therapy* 35, no. 7: 633–649. https://doi.org/10.1016/S0005-7967(97)00018-1.

———. 1999. “Evolutionary versus Prototype Analyses of the Concept of Disorder.” *Journal of Abnormal Psychology* 108, no. 3: 374–399. https://doi.org/10.1037//0021-843x.108.3.374.

———. 2000. “Spandrels, Vestigial Organs, and Such: Reply to Murphy and Woolfolk’s ‘The Harmful Dysfunction Analysis of Mental Disorder’.” *Philosophy, Psychiatry & Psychology* 7, no. 4: 253–269. http://dx.doi.org/10.1353/ppp.2000.0040.

———. 2007. “The Concept of Mental Disorder: Diagnostic Implications of the Harmful Dysfunction Analysis.” *World Psychiatry* 6, no. 3: 149–156. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2174594/pdf/wpa060149.pdf.

———. 2014. “The Biostatistical Theory versus the Harmful Dysfunction Analysis, Part 1: Is Part-Dysfunction a Sufficient Condition for Medical Disorder?” *Journal of Medicine and Philosophy* 39, no. 6: 648–682. https://doi.org/10.1093/jmp/jhu038.

———. 2015. “DSM-5, Psychiatric Epidemiology and the False Positives Problem.” *Epidemiology and Psychiatric Sciences* 24, no. 3: 188–196. https://doi.org/10.1017/S2045796015000116.

———. 2016. “Diagnostic Issues and Controversies in DSM-5: Return of the False Positives Problem.” Annual Review of Clinical Psychology 12 :105–132. https://doi.org/10.1146/annurev-clinpsy-032814-112800.

———. 2021. “Quinian Qualms, or Does Psychiatry Really Need the Harmful Dysfunction Analysis? Reply to Harold Kincaid.” In *Defining Mental Disorder: Jerome Wakefield and His Critics,* edited by Luc Faucher and Denis Forest. Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/9949.003.0011.