13 Function and Dysfunction

Dominic Murphy

Introduction

Few ideas in the philosophy of psychiatry have been discussed as widely as Jerry Wakefield's harmful dysfunction analysis (HDA) of mental disorder in the twenty years or so since its original promulgation (Wakefield 1992a, 1992b, 1993). This is unsurprising, since it is a tour de force of simple, elegant analysis. And it does seem to get the basic picture right; there is something compelling about the idea that the attribution of mental illness rests on both a value judgment and a belief that something is wrong with the inner systems of a sufferer. On closer inspection, though, I think that the HDA turns out to be unconvincing when it comes to the details. The HDA's account of function is not the right one for the job. The philosophical literature on functions has uncovered many nuances that Wakefield misses, and the evolutionary analysis of functions makes a number of commitments that do not seem to be part of the general tendency to attribute mental illness that Wakefield relies on as evidence for his account. It also seems that medicine and other branches of experimental biology rely on a different account of function than the evolutionary one. I shall lay out Wakefield's analysis and, in doing so, point to some of its attractions. I will take a while to describe the HDA as I understand it. This is partly because I think that some criticisms of the HDA are unfounded and I want to dissociate myself from them, but more important because some of those criticisms would, if correct, tell against a broader naturalistic family of analyses of which the HDA is a member. I think that once the HDA is laid out and its attractions manifest, we can see that there are further options for naturalistic analyses.

Although I will be critical of Wakefield's HDA in what follows, I do wish to be constructive and exploratory, as well. Wakefield has for the most part been clear and concise in advancing his theory and bitten some bullets in defending it. This lets us see what is at stake, and one important result is that Wakefield has opened up ways to develop a family of views that are not simply normative. In what follows, then, I will suggest ways in which the details of the HDA might be changed but remain in keeping with its basic philosophical orientation.

I. The Components of Wakefield's View

1.1 The HDA as a Two-Stage View

I class Wakefield's HDA as a two-stage account of mental disorder, in the tradition of Boorse (1975, 1976). Two-stage theorists hold that there are two individually necessary and jointly sufficient conditions for disorder. First, there is a biological dysfunction. Wakefield's innovation was to see this as specifically a failure by a bodily system to perform the naturally selected function that explained the system's replication in past generations: "the failure of a mechanism in the person to perform a natural function for which the mechanism was designed by natural selection" (Wakefield 1993, 165).

Second, the dysfunction must result in harm to the individual. "Harm" is generally recognized to be a normative notion, and Wakefield thinks we follow a simple rule when judging that someone is harmed. It is judged by prevailing social norms: "defined by social values and meanings" (1993, 373). The fact that this is a simple rule does not mean that it is a simple matter to tell where and how it applies. Whether somebody is harmed may be difficult to assess, and although judgments can be uncontroversial (a terminal disease or a serious injury is obviously harmful), they often won't be. But although assessing harm is often difficult or controversial, the rule is simple: harm is assessed relative to the prevailing norms of the society, not the views of the individual concerned, who may not feel as though they are badly off at all.

In sum, we have the two components of the HDA (Wakefield 2006, 157), which state that for a condition to count as a mental disorder: (1) it is negative or harmful according to cultural values, and (2) it is caused by a dysfunction (i.e., by a failure of some psychological mechanism to perform a natural function for which it was evolutionarily designed).

The mix is part normative and part objective. Clause (1) tells us that what counts as harmful will differ across times and cultures, but clause (2) tells us that dysfunction will not. Our psychology, like our physiology more generally, is made up of numerous mechanisms that combine to cause normal behavior. Each of these has the job that it was selected for, and if it does not work as selected, it is dysfunctional. On some occasions, that dysfunction causes some phenomenon—physiological, mental, behavioral—that is judged to be harmful by the wider group that the person with the dysfunction belongs to. At other times, the dysfunction might not be a source of harm. In theory, a dysfunction could make you better off. For example, if your liver metabolism departed from its historically adaptive range in a way that enabled you to drink all you want without suffering ill effects, people might not think that you were harmed, at least in the circles I move in. I think it is clear that a view of this type—a two-stage view—has many attractions, which perhaps come out if we compare it to its main rivals. Doing so will also raise some important issues that will detain us later. I will go over these attractions and questions now.

1.2 The HDA and Normative Views

Two-stage views sit between purely normative views and those that deny that our disease concepts employ any human norms at all. On one hand, a purely normative view would argue that an attribution of mental illness is solely a value judgment: it merely reflects the way we evaluate people with respect to prevailing social norms. A view like this faces the great problem that we distinguish many forms of departures from normality. Sometimes we judge people to be ill or mentally disordered, but we can also judge them to be deviant in a host of other ways. They might be criminal, or eccentric, or immoral, or indeed they might depart from normal forms of behavior in ways that we prize, in which case we call them geniuses. What, on the purely normative view of mental disorder, makes the difference in these cases? Why do we call some behavior eccentric or criminal and other behavior disordered? One big advantage of a two-stage account of mental disorder is the answer it gives to this question. We call people disordered or diseased when we think that there is something wrong "under the hood." Some part of them is not working the way that a normal component of human being naturally works, and it is responsible for the salient features of their behavior. I take it that this is the basic intuition Wakefield is working with. We respond to the mentally ill as if there is something wrong under the hood—that is the dysfunction, and we take the salient harm to be evidence for it.

Another way of asserting the purely normative view is to insist that calling something a dysfunction expresses a stance toward it, perhaps based on what we take to be the best interests or purposes of the owner of the system. I think Wakefield is quite correct to push back against claims of this kind (e.g., in Wakefield 2009, 92–93): it is one thing to identify a system as relevant for its owner but another to assert that its function is just normative. If you are training someone to be an athlete or an opera singer, it might pay to enhance their lung capacity. In that sense, their respiratory system is a site of our joint interests, but the function of that system is still perfectly objective. What a sophisticated proponent of the purely normative view should do is argue that disease concepts have a different structure altogether. Cooper (2007) and Murphy (2006) have drawn an analogy between the concept of mental disorder and that of weed. Weeds are not a natural kind. We can perhaps say that a weed is a fast-growing species that negatively impacts on economically valuable crops, usually through competition for nutrients, sunlight, and space. What fixes the extension of "weed" (and similar concepts like "vermin" or "precious metal") is a set of contingent human interests that can change over time. Suppose that determining that a condition is a disorder is like determining that a plant is a weed. The judgment is determined by normative considerations that we have already made. But nonetheless, there is real, explanatory mind-independent knowledge to be had about each sort of "weed."

A skeptic about the objectivity of mental disorder could exploit this point. She could admit that there are correlations between psychologically salient behavior and

the performance of underlying systems. Such a skeptic might also argue that although there is perfectly objective knowledge to be had about those systems, they only count as disordered because of our prior decision to insist that they are disordered—anything that produced that behavior would be called dysfunctional, because it is productive of mental illness. Such a skeptical view would also combine a claim about judgments of disorder with a claim about the objective nature of underlying systems. It could even agree that commonsense attributions of mental illness involve detection of systems that strike us as broken "under the hood." It would, however, deny that there is any sense to calling such systems dysfunctional. So a two-stage theory like the HDA needs a way to show that a system is dysfunctional that meets this skeptical challenge.

The last point raised by a purely normative analysis of the concept of mental disorder is that of relativity. If being mentally ill is just a matter of how it strikes the other members of one's society, then it seems that whether you are mentally ill depends on where and when you live. The history of psychiatry—indeed, of the sciences of the mind in general—is full of episodes in which some disfavored group has been condemned as pathological in the light of considerations that strike us as fraudulent. Homosexuals, notoriously, were diagnosed as mentally ill until the revision of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*) in the 1970s. The HDA lets us say that such judgments are incorrect unless they involve the detection of some underlying disorder, whereas a purely normative view seems to have no resources to avoid relativism.

So far in this section, I have contrasted the HDA with views that insist on the purely normative character of the concept of mental disorders. The other option would be a completely nonnormative view, insisting that judgments of harm are just as objective as those of dysfunction. A two-stage view incorporates values judgments and, in so doing, fits both medical practice and ordinary thought better, because it does seem undeniable that we argue about whether something is a disease based on whether we think it makes a life go less well, and the relevant considerations are indeed a matter of socially constituted meanings.

The story so far is this: the HDA is one of a family of two-stage analyses of the concept of disorder that combine normative evaluations of harm with criteria of dysfunction that aim to be objective and scientifically grounded. In the second part of the chapter, I will discuss whether, as Wakefield contends, the analysis of function that this account needs is an evolutionary one. But first, I will mention one last issue in the characterization of the HDA. Whose concept is being analyzed?

1.3 Whose Concept?

Conceptual analysis in philosophy is often undertaken with the aim of regimenting nontheoretical talk. We might pick out a term that plays a role in commonsense description of the world and try to define it. Many philosophers of science, though, are interested in how terms work in theoretical contexts; concepts exist in unfolding

scientific projects, and we want to understand how they work. When we do this, we often uncover differences or ambiguities in scientific practice. Philosophers of biology, for example, have largely come to agree that there is no unified concept of the gene throughout biology (Griffiths and Stotz 2013). Rather, different areas of biology, or even the same scientist at different times, will use the gene concept to pick out different aspects of the underlying genetic and developmental processes. I think *function* is scientifically polysemous in just this way and will argue for that claim in the next section. Right now, I want to ask what sort of project Wakefield is engaged in.

Wakefield typically argues that his project involves the analysis of *mental disorder* in scientific and clinical contexts. However, Wakefield is also happy to argue that some diagnoses are in error because they ignore the lessons of common sense.

Horwitz and Wakefield (2007), for example, argue that the *DSM-IV* fails to respect common sense or previous psychiatric consensus about depression, and as a result is diagnosing many people as depressed when they are just normally miserable. Since classical antiquity, Western thought has recognized a condition of melancholy, as people slip into depression without any observable, proportionate cause. This is to be distinguished from ordinary understandable sadness that people suffer when they are visited by life's misfortunes. The tradition sees pathology only in the former case, insisting that "pathological depression is an exaggerated form of a normal human emotional response" (Horwitz and Wakefield 2007, 71). They conclude that the concept of depression defined by the contemporary diagnostic syndrome represents a major conceptual break with both past psychiatry and commonsense thought about human nature. This has led to needless alarmism about an epidemic of depression and caused misfortune for many individuals diagnosed in error.

It is correct that the *DSM* concept of depression, on the face of it, lumps together different psychological and behavioral types in the same category because of observable similarities, and it is quite true that observable similarities may nonetheless reflect diverse etiologies. The response (Kendler et al. 2008; Zisook and Kendler 2007) is to argue that the populations are too similar on the relevant measures to justify treating them as different. This argument was made with great force in the run-up to the *DSM-5*, when various parties debated whether it was appropriate (as in *DSM-IV*) to exempt the bereaved from a diagnosis of depression or to, in effect, treat grieving as form of depression.

Wakefield has weighed in on the statistical issues involved in this debate (Wakefield et al. 2007), but it is the conceptual argument I want to focus on here. If we are trying to capture a scientific concept, on what grounds can we argue that it should be criticized for departing from traditional conceptions? A chemist who was told that traditional usage did not regard objects as made of atoms would be unmoved. So, one possibility is that Wakefield is just being inconsistent, appealing to scientific practice in some cases and traditional or commonsense views in others. A more charitable reading

would be that Wakefield thinks that the underlying logic of psychiatric practice accords with the HDA but that sometimes psychiatrists depart from the logic of their own projects. Nonetheless, this does raise the question of how revisionist the HDA might turn out to be. It is unusual for scientific developments to vindicate traditional modes of thought, such as the venerable distinction between melancholy and depression, but Wakefield's view, as has been pointed out (Kingma 2013; Murphy and Woolfolk 2000a, 2000b), looks to be very revisionist for independent reasons, connected to the HDA's requirement that dysfunction is necessarily a property of evolved systems.

1.4 Revisionism

Many critics have argued that the existence of evolutionary by-products is a problem for Wakefield's account (Kingma [2013, 375–379] has an excellent review). Wakefield himself notes that acalculia, dyslexia, and amusia are disorders that do not involve selected effects but side effects of other systems (2011, 169). On the face of it, the HDA would have to rule these conditions out, but Wakefield has a reply. He argues that in dyslexia, the problem must arise because of a dysfunction in some other system that does have a selection history.

Now, as Kingma (2013, 376) notes, it is quite possible for a linked trait to be abnormal for independent reasons that have nothing to do with a dysfunction in the system it is linked to. Wakefield (2011, 170) argues that students of reading disorders assume that there is a dysfunction involved in the condition, which they do not do for cases of ordinary illiteracy. This is correct, but it only meets the objection if those researchers are committed to an evolutionary account of function, which is the very point at issue. The fact that they take for granted the existence of a dysfunction (or some other proximate cause of the condition) does not show that they take for granted the existence of an evolutionary dysfunction. Wakefield thinks that he can rule this out because there is no viable concept of dysfunction other than an evolutionary one, but that is not correct, as the next section will show.

Kingma argues that Wakefield's position is potentially very revisionist indeed: if a large number of disorders turn out not to be linked to selected effects but to be the product of other processes, such as developmental or genetic linkage, then he will have to say that they are not really disorders. Wakefield seems to think that this is not a live possibility and that one can always postulate the existence of a selected function that is linked to the condition. As Kingma (2013) and Nordenfeldt (2003) argue, this seems grossly ad hoc: one can indeed always postulate such a dysfunction, however remote the causal chain may be, but this just makes the HDA immune to counterexample by stipulation. Whenever the apparent cause of a disorder seems to be something other than a dysfunctional selected system, Wakefield can just insist that there must be one somewhere.

Suppose the view is revisionist, and we have to revise our thinking about mental disorder. Is that so bad? Well, psychiatric concepts, like some others, are impure: although

they involve scientific criteria, they also speak to urgent practical needs. Part of the appeal of the HDA is its impurity, since it incorporates both scientific assessment and normative, cultural assessment. If it turned out that a lot of what we consider to be mental disorders are not disorders at all, we would still worry about the people who fall under those diagnoses, and we would still want concepts that differentiate them from others. And those concepts would need to direct our attention to the causes of the conditions, because we would want to be able to intervene in order to improve their lot. Certainly, some human traits have come to be seen as normal rather than disordered, and others have gone the other way. But widespread revisionism seems unattractive, and if it can be avoided, I take that to be a point in favor of an analysis.

1.5 Some Desiderata

I have argued that Wakefield's HDA is an instance of a two-stage view that combines a normative component and a naturalistic understanding of function. The question now before us is whether his evolutionary understanding of function is the correct one. I extract three desiderata from the discussion so far that can serve as constraints on the account of function we need. First, it should be able to make precise the idea that attributions of mental illness rely on something being objectively wrong under the hood of a human being. Second, it should capture relevant scientific practice. Third, it should not be too revisionist but make sense of the idea that some conditions may have causes that cannot be shown to depend on failures of selected effects. In the next section, I will argue that there are accounts of function available that do a better job of meeting these constraints than the evolutionary account that Wakefield prefers. Therefore, there are two-stage analyses of the concept of mental illness that are superior to the HDA while still retaining its benefits.

II. The Concept of Function[1]

As I said in the last section, the HDA is based around the fundamental intuition that our ascriptions of mental disorder reflect the view that something is wrong under the hood of a human being. It assumes that what is wrong is a dysfunction in a system whose normal function can be objectively discovered. Wakefield claims (2011, 144) that Darwin discovered the nature of biological function, just as atomic chemistry discovered the nature of water. Therefore, the Darwinian understanding is what people have always referred to when talking of natural functions. Hence, the HDA (or any two-stage theory) must appeal to a Darwinian notion of function to make sense of attributions of dysfunction in medicine, psychiatry, and ordinary unscientific talk, for "function" identifies a set of beneficial capacities of living things, and Darwin has told us what those are. When there is a failure of natural order, it is of the order imposed by natural selection.

If it were correct that a Darwinian account of function was the only scientifically respectable adumbration of that concept, then Wakefield's position would be unimpeachable. But it is not correct. In contemporary philosophy of biology, analyses of function derive from two important papers. Wright (1973) argued that ascriptions of function to a structure are causal-historical, where function depends on a prior selection process. Wright is sometimes taken to have adopted an evolutionary account of function that relies on the notion of natural selection, but he did not. It was Millikan (1984) and Neander (1991) who developed Wright's account into an explicitly evolutionary one. Wakefield's understanding of function is squarely in this tradition.

Cummins (1975) was the other key paper. Cummins's concept of function was not historical but dispositional. He understood the function of an entity to be the contribution it makes to "an adequate analysis" of the capacity of the overall system that includes it. According to Cummins, a component may have a function even if the component was not "designed," and therefore, parts with no selection history can be ascribed a function. Wakefield argues that Cummins's account was introduced to capture functional explanation in the philosophy of mind and is not relevant to biology (2011, 149). However, philosophers of biology (including Cummins himself) have elaborated the analysis, just as Millikan and Neander did with Wright's historical account, and they have done so in order to capture the important role that nonhistorical concepts of function play in many areas of biology. In doing so, they have shown how this concept of function fits into the mechanistic explanations that are common in biomedicine. In this section, I will introduce the selectionist analysis of function that Wakefield prefers, then the alternative causal-mechanical analysis, and argue that the latter better fits medical practice, is epistemically less committed, and is at least as good a fit for common sense.

2.1 The Selectionist View

Evolutionary views of function involve causal-historical explanations of traits that I will call *selectionist*. The heart is a standard example. Millikan (1984) said that *the heart is a pump because it is the heart's pumping that contributes to the successful reproduction of organisms with hearts*: if x is a member of a biological category, it is not because of "the actual constitution, powers, or dispositions" of x, but because of the "proper function" of x (Millikan 1984, 17). X's proper function depends on the history of x's lineage, which explains x's being supposed to do whatever it does. Neander (1991, 180) agrees that a biological part is only identifiable in terms of its proper function.

The point is quite subtle because the relevant history consists of correlations obtained between ancestors of x having a certain character and their having been able to perform x's function. So the structure of a heart explains why it pumps, but it does not count as a heart (or a pump) in virtue of having that structure.

The selectionist account of function seems to offer two big benefits. First, it promises to give a definite specification of the function of an organic system and hence

a clear criterion for calling it dysfunctional. Second, it seems to offer a scientifically unproblematic way to say what a system ought to be like. If you are worried about the accusation that function talk is normative, you can embrace natural selection. Teleological notions are commonly associated with the pre-Darwinian view that the biological realm provides evidence of conscious design by a supernatural creator. The point about selectionist concepts of function is that they assuage this metaphysical concern by showing how norms are part of nature. I am not going to suggest that there is something wrong with the Darwinian picture of natural order. But I will suggest that selectionist concepts of function and dysfunction are a poor bet for psychiatry. I will now introduce a rival account of function, but before developing it philosophically, I will try to motivate it: the point is not just to argue that there is a philosophical option that Wakefield has not considered but to argue that this unconsidered option fits the relevant sciences better. So I will start with the science.

2.2 Why We Need Other Concepts of Function

Wakefield argues that selectionist accounts of function solve the "essential explanatory puzzle posed by function attributions within biology" (2011, 149), namely, how can there be apparently designed systems in a world devoid of purpose. However, this metaphysical question, despite its importance, is not the only context within which functional talk appears in the sciences. As Amundson and Lauder (1998, 227) put it, "Philosophers' special interests in purposive concepts can lead to the neglect of many crucial but non-purposive concepts in the science of biology." Amundson and Lauder maintain that a selectionist analysis of function fits evolutionary biology, but they argue that a different concept of function is used in comparative and functional anatomy. They contend that this alternative concept of function is well captured by a Cummins-style account, which they call a causal-role analysis. In contrast to Millikan and Neander, they point out that it is entirely possible to identify anatomical units by anatomical considerations, regardless of proper function. They also argue for the ineliminability of a causal-role analysis of function, on the grounds that an anatomical unit can have a function even in the absence of a selectionist history: the fact that a biological system has a selective history does not imply that all of its components have a distinct selective history that makes them what they are. Their functions, in the sense of their causal contribution to the overall system, are independently identifiable on morphological and physiological grounds, regardless of history.

The biomedical sciences routinely try to work out what a system contributes to the overall functioning of the organism. In doing so, they typically do not try to establish that a biological component has a selectionist function. For example, take Hubel and Wiesel's famous program of mapping the receptive fields of cells in the visual cortex and then establishing further visual information-processing channels in the brain. That program, and the research on the neurobiology of vision inspired by it, depended on a

set of engineering assumptions about the way the brain is organized to process information. It did not test assumptions about the selective advantage and history of the components of the visual brain. Hubel and Wiesel never sought to show that the cats whose visual system they interfered with produced fewer descendants than other cats.

It may be that the facts uncovered in physiology are evidence for evolutionary relationships, and of course, all biological systems have an evolutionary history, but when we determine what normal function is, in medicine, we do not even try to establish what that history is: selectionist, or broader historical, considerations do not arise. It is the mechanistic relations between parts of the system that matter.

Schaffner (1993) argued that although medicine might use teleological talk in its attempts to develop mechanistic explanations, that talk is just heuristic. It focuses our attention on entities that are useful to the organism. Schaffner suggested that as we learn more about the role a structure plays in the overall functioning of an organism, the need for evolutionary functional ascriptions drops out. It is replaced by the vocabulary of mechanistic explanation: the causal relationship of parts that jointly produce phenomena of explanatory interest. Functional explanations that draw on evolutionary considerations are, he claimed, "necessary, but empirically weak to the point of becoming almost metaphysical" (Schaffner 1993, 389–390).

Normal biomedical ascription of function to a system makes no claims about selective history. It requires only that we can identify the role played by a system in the overall economy of the organism. How is dysfunction determined? By the use of a biomedical concept of normality that is an idealized description of a component of a biological system in an unperturbed state. It does not rest on the failure of a biological part to replicate as its ancestors did, or to reduce overall fitness, but by its failure to be close enough to the causal contribution of the analogous part in the idealized overall system.

Wachbroit (1994, 588) argues persuasively that when medicine or physiology says that an organ is "normal," the relevant conception of normality "is similar to the role pure states or ideal entities play in physical theories." Such an idealization represents actual organs or systems in unperturbed states (see also Ereshefsky 2009). To understand a real case, we add information to develop a model that resembles actual hearts (Wachbroit 1994, 589). For instance, Gross (1921) was able to establish post mortem that anastomotic communication between main arteries increases over a typical life span, thereby establishing that we need to model younger and older hearts differently. The point of such idealizations is not to represent the statistically average heart but to describe hearts in a way that allows departures from the ideal to be recognized and to serve as template from which more realistic models can be built.[2] In general, physiological theories are families of such idealizations, and bodily systems are understood as functional parts of larger systems, typed unhistorically.

Insofar as psychiatry is a branch of medicine, the concept of function it needs resembles those of physiology and biomedicine. Evolutionary considerations are just beside the point. Wakefield argues that the selective account of function is one that should

be embraced by the philosophy of medicine because it is the working assumption of medical scientists. That is just not so: most parts of medicine and biology, including the areas closes to psychiatry, use a mechanistic, ahistorical account of function. The concepts of psychiatry should be continuous with those of medicine and physiology more generally. The life sciences ask all sorts of questions, but the questions that medicine asks are not those that a selectionist account of function can answer. Most biomedical research is based on establishing the components, as well as the functional relations between components, in biological systems. It is not aimed at uncovering evolutionary relationships. Health and fitness are different concepts, with different functional analyses. As Gluckman et al. (2009, 4–8) argue in their textbook on evolutionary medicine, for example, medicine is about health, and evolutionary biology is about fitness, and the latter does not provide a definition of disease.

At the end of section II, I suggested that there were three constraints on a satisfactory account of function from the point of view of a two-stage analysis of mental disorders such as the HDA. The second of these was that it should capture biomedical practice. The selectionist account of function does not do this as well as the rival, ahistorical account, which from now on I will call the *systemic capacity* account. That leaves two other constraints: whether the account can provide an objective articulation of the intuition that there is something inner—under the hood—that is wrong with the mentally ill and whether it will deal with cases of disorders that do not seem to depend on failures of selected effects. I think that the systemic capacity account does well on both counts, and I will now try to show why. I have argued for the ubiquity of an alternative analysis of function in "experimental biology" (Weber 2005b) and medicine but not provided many details. In the last part of the chapter, I will sketch the account I prefer and argue for its virtues.

2.3 The Systemic Capacity Account

Cummins (1975) introduced his account of function in the context of explaining how the overall capacity of system—its ability to do something—depends on the subcapacities that interact to produce it. A components function in a system is whatever it does that contributes to the overall capacity of the system that contains it. As it stands, this view does not tell us why the entity with the component's function is there in the first place, which is what the historical account was designed to do. However, as Cummins (2002) points out, a selectionist account of function does not say why the entity is there in the first place either. Selection accounts for the spread of a trait, not its origin. In order to outcompete its variant, a system must exist, and the selectionist story does not tell us why it exists.

The bigger problem with Cummins's analysis is that, notoriously, attributions of function on this account are interest relative: if you are interested in the heart's contribution to the circulation of the blood, you can analyze it one way, and if you are interested in its contribution to the sum total of the noises the body makes, you can treat it as a "lub-dup" generator. In both cases, it counts as a part of the overall system you wish to

decompose. Cummins thought that the overall system depended on the interests of the investigator. In that sense, he was not really trying to articulate a naturalistic concept of function that captures biological practice but to understand the logic of functional analysis. This is, however, a problem for such a view with respect to the first of my three constraints: we want an objective characterization of dysfunction rather than a mere acknowledgment that we decompose a system into its contributory parts. If that were all that the account gave us, it would threaten the objectivity that is appealing about two-stage views a like the HDA. It would be possible to argue like the skeptic I imagined in section 2.2. The skeptic contends that our analysis of what biological systems do is driven by a prior judgment that they are disordered. It is not independent of that judgment *but is a scientific analysis of what we have decided to call a disorder on evaluative grounds*. To rebut the skeptical point, we need a way of characterizing biological systems that sees them as mind-independent components of nature, not just reflections of human concerns.

As Amundson and Lauder note (1998, 237), even within unambiguously scientific contexts, it is possible to cook up "whimsical Cummins functions" such as Neander's (1991) example of the function of geological plate movements in tectonic systems. What the whimsical examples trade on is the absence in Cummins's account of any commitment to the overall goal of the system; there is nothing that the tectonic system is for, so attributing functions to its parts looks weird. However, Amundson and Lauder argue that these examples do "not apply to the real world of scientific practise." They argue that in fields like comparative anatomy, one finds anatomical, rather than purposive or functional, characterizations of living systems with causal relations among their parts that a Cummins-style account is needed to deal with.

I have some sympathy with the idea that the whimsical objections to Cummins's analysis are scientifically irrelevant. Nonetheless, it would be good to have some principled way of identifying natural systems. Recent attempts to do this attempt to identify some metaphysical relation that holds natural systems together.

Weber (2005a) argues for *coherence*. A coherent system is one that displays a complex network of capacities with contributory relations among them, so that capacities contribute to other capacities that contribute to other capacities. In his example (193), ion channels in nerve membranes "regulate ion permeability because this capacity is part of an account of the nervous system's capacity to process information. Therefore, it is a function of nervous membranes to fire action potentials. Furthermore, the nervous system's capacity to process information is part of an analytic account of the organism's capacity to locate food and sexual partners" and so on. Organisms are webs of integrated explanatory relations, and respecting these webs provides a constraint on individuating systems that makes them not just choices of an investigator but genuine parts of nature.

A related account tied to a fuller account of explanation is Craver's (2007) development of Cummins-style functional analysis into mechanistic explanation. Craver begins his book by discussing the mechanism by which neurotransmitters are released

(4–6). This involves finding answers to questions such as the following: why does depolarization of an axon terminal lead to neurotransmitter release, and why are neurotransmitters released in quanta? Answers identify anatomical entities, such as specific types of calcium and various intracellular molecules, and show how they interact with each other to give rise to the explanandum. This picture applies to mental disorders just as it does to the activities of the normal mind: they depend on the interaction of biological systems that manifest a given phenomenon. An explanation with these features is mechanistic: it appeals not to natural laws but to the interaction of natural systems.

In recent years, philosophers have stressed the way in which explanation in many sciences, above all the biological and cognitive, depends on finding the parts within a system whose interactive structures and activities explain the phenomena produced by the system. Philosophers disagree over exactly how to characterize mechanisms, but it is agreed that mechanisms comprise (1) component parts that (2) do things. Strife arises over how to see the activities of the parts. Are they also primitive constituents of a mechanism or just activities of the constituent components (for full references and a review, see Tabery 2004). A mechanistic explanation shows how these parts and what they do give rise to the phenomenon we want to explain.

I take it that the affinity with a Cummins-style account of function is clear. Craver (2007, 161) asks, "How must Cummins-style functional analysis be restricted to provide a normatively adequate of mechanistic explanation" (i.e., one that displays the properties that good explanations have)? His answer, which is in the spirit of Weber's story, is that relations between subcomponents of a system must be constitutive. A mechanism has parts that hang together and jointly compose a system, and the causal relations among them must be genuine ones, rather than mere correlations. I will not go over Craver's full account, which is thorough and elaborate. The point is that philosophers of biology have worked to develop a Cummins-style analysis into a genuine causal-explanatory account that fits into a wider picture of how organisms fit together and how mechanistic explanation reveals the explanatory relations among parts of biological systems. Wakefield's claim (2011, 149) that the Cummins picture *"does not attempt to elucidate how functional relations are explanatory"* of causal relations in biology is true of the original account. But it overlooks the extensive work done in recent years by philosophers of biology who have developed it into a causal-explanatory account.

2.4 The Scoreboard

At the end of section II, I listed three constraints on an account of function suitable for a two-stage theory of mental disorder like the HDA. It is time to take stock. I have already shown that when it comes to capturing biological practice, the systemic capacity view is superior, because it accords with the understanding of function used in the parts of biology and medicine that are most relevant to psychiatry, rather than evolutionary biology. What about the other two constraints that the account of function should meet? One of them was the extent to which an account of function can provide

an objective understanding of what we mean when we attribute a dysfunction. I have argued, although not shown in detail, that a systemic capacity account can comport with an objective understanding of natural systems. Exactly what ordinary attribution of dysfunction means is unclear. However, there is no reason to suppose that it must be articulated in an evolutionary fashion—certainly we don't check whether someone has low biological fitness before we call them disordered. Perhaps we should say here that no view has a clear advantage. My own hunch is that there is no one view that is the best candidate for the scientific articulation of every kind of selection talk, including the ways in which we attribute dysfunction. But there is no reason to think that the systemic capacity view cannot serve as well as the selectionist view when it comes to the relevant contexts. There is no reason to think that comparative anatomy or neurophysiology is less objective than evolutionary biology, not that they are lesser candidates than evolutionary biology for the role of capturing the sorts of judgment Wakefield is interested in.

Last, I suggested that a view might want to escape the apparent revisionism that lurks in Wakefield's commitment to the necessary existence of an evolutionary dysfunction. Here the systemic capacity view wins easily. The systemic capacity view can ask what in the individual is contributing to the salient behavior without worrying about history. Wakefield stipulates that there must be some evolved system somewhere that has departed from its historical design and is connected to the current problem. In contrast, the systemic capacity view just asks *what the underlying system is that is misbehaving*. This is what I mean by calling it less epistemically committed. It has one less bet to make. We both agree that researchers in dyslexia, say, are looking for dysfunctions. Wakefield assumes that they must be concerned with the history of the systems they scrutinize. The systemic capacity view just says that they must be concerned with what the systems are like. Even if a system has no selective history, it can still fail to exhibit its normal capacity and be treated as dysfunction. The systemic capacity view can deal with these cases, and the selectionist view cannot.

So the systemic capacity view is a winner on at least two counts out of three, and one is at best a stalemate. I suggest that this means that the "dysfunction" part of the HDA needs to be rethought and brought into line with contemporary biology and philosophy of biology. The HDA remains a very elegant and attractive analysis, but its concept of function will not do.

Notes

1. Parts of this section draw on Roe and Murphy (2011).

2. Notice that some degree of idealization is required by a selectionist account too; one indisputable result of Darwin's work is the demonstration that variation is ubiquitous in nature. Any determination of the normal range of function of a biological system requires some idealization to cope with that variation.

References

Amundson, R., and G. Lauder. 1998. Function without purpose. In *The Philosophy of Biology,* D. L. Hull and M. Ruse (eds.), 227–257. Oxford University Press.

Boorse, C. 1975. On the distinction between disease and illness. *Philosophy and Public Affairs* 5: 49–68.

Boorse, C. 1976. What a theory of mental health should be. *Journal for the Theory of Social Behavior* 6: 61–84.

Cooper, R. 2007. *Psychiatry and the Philosophy of Science*. Acumen.

Craver, C. F. 2007. *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford University Press.

Cummins, R. 1975. Functional analysis. *Journal of Philosophy* 72: 741–764.

Cummins, R. 2002. Neo-teleology. In *Functions: New Essays in the Philosophy of Psychology and Biology,* R. Cummins, A. Ariew, and M. Perlman (eds.), 157–173. Oxford University Press.

Ereshefsky, M. 2009. Defining "health" and "disease." *Studies in the History and Philosophy of Biology and Biomedical Sciences* 40(3): 221–227.

Gluckman, P., A. Beedle, and M. Hanson. 2009. *Principles of Evolutionary Medicine*. Oxford University Press.

Griffiths, P., and K. Stotz. 2013. *Genetics and Philosophy: An Introduction*. Cambridge University Press.

Gross, L. 1921. *The Blood Supply to the Heart in Its Anatomical and Clinical Aspects*. New York: Hoeber

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

Kendler, K. S., J. Myers, and S. Zisook. 2008. Does bereavement-related major depression differ from major depression associated with other stressful life events? *American Journal of Psychiatry* 165(11): 1449–1455.

Kingma, E. 2013. Naturalist accounts of mental disorder. In *The Oxford Handbook of Philosophy and Psychiatry,* K. W. M. Fulford, M. Davies, R. G. T. Gipps, G. Graham, J. Z. Sadler, G. Stanghellini, and T. Thornton (eds.), 363–384. Oxford University Press.

Millikan, R. G. 1984. *Language, Thought and Other Biological Categories*. MIT Press.

Murphy, D. 2006. *Psychiatry in the Scientific Image*. MIT Press.

Murphy, D., and R. L. Woolfolk. 2000a. The harmful dysfunction analysis of mental disorder. *Philosophy, Psychiatry, and Psychology* 7(4): 241–252.

Murphy, D., and R. L. Woolfolk. 2000b. Conceptual analysis and scientific understanding. *Philosophy, Psychiatry, & Psychology* 7(4): 271–293.

Neander, K. 1991. Functions as selected effects: The conceptual analyst's defense. *Philosophy of Science* 58(2): 168–184.

Nordenfeldt, L. 2003. On the evolutionary concept of health: Health as natural function. In *Dimensions of Health and Health Promotion,* L. Nordenfeldt and P.-E. Liss (eds.), 37–53. Rodopi.

Roe, K., and D. Murphy. 2011. Function, dysfunction, adaptation? In *Maladapted Minds,* P. Adriaens and A. Block (eds.), 216–237. Oxford University Press.

Schaffner, K. F. 1993. *Discovery and Explanation in Biology and Medicine*. University of Chicago Press.

Tabery, J. 2004. Synthesizing activities and interactions in the concept of a "mechanism." *Philosophy of Science* 71: 1–15.

Wachbroit, R. 1994. Normality as a biological concept. *Philosophy of Science* 61: 579–591.

Wakefield, J. C. 1992a. The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist* 47(3): 373–388.

Wakefield, J. C. 1992b. Disorder as harmful dysfunction: A conceptual critique of *DSM-III-R*'s definition of mental disorder. *Psychological Review* 99(2): 232–247.

Wakefield, J. C. 1993. Limits of operationalization: A critique of Spitzer and Endicott's (1978) proposed operational criteria for mental disorder. *Journal of Abnormal Psychology* 102(1): 160–172.

Wakefield, J. C. 1999. Disorder as a black box essentialist concept. *Journal of Abnormal Psychology* 108: 465–472.

Wakefield, J. C. 2006. Personality disorder as harmful dysfunction: *DSM*'s cultural deviance requirement reconsidered. *Journal of Personality Disorders* 20: 157–169.

Wakefield, J. C. 2009. Mental disorder and moral responsibility: Disorders of personhood as harmful dysfunctions, with special reference to alcoholism. *Philosophy, Psychiatry, and Psychology* 16: 91–99.

Wakefield, J. C. 2011. Functional explanation: Aristotle, Lucretius, Darwin. In *Maladapted Minds*, P. Adriaens and A. Block (eds.), 143–172. Oxford University Press.

Wakefield, J. C., M. F. Schmitz, M. B. First, and A. V. Horwitz. 2007. Extending the bereavement exclusion for major depression to other losses: Evidence from the National Comorbidity Survey. *Archives of General Psychiatry* 64: 433–440.

Weber, M. 2005a. Holism, coherence and the dispositional concept of functions. *Annals of the History and Philosophy of Biology* 10: 189–201.

Weber, M. 2005b. *Philosophy of Experimental Biology*. Cambridge University Press.

Wright, L. 1973. Functions. *Philosophical Review* 82(2): 139–168.

Zisook, S., and K. S. Kendler. 2007. Is bereavement-related depression different than non-bereavement-related depression? *Psychological Medicine* 37(6): 779–794.

Defining Mental Disorder

Jerome Wakefield and His Critics

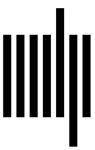
By: Harold Kincaid, Peter Zachar, Dominic Murphy, Justin Garson, Philip Gerrans, Rachel Cooper, Steeves Demazeux, Leen De Vreese, Maël Lemoine, Tim Thornton, Andreas De Block, Jonathan Sholl

Edited by: Luc Faucher, Denis Forest

Citation:

Defining Mental Disorder: Jerome Wakefield and His Critics
By: Harold Kincaid, Peter Zachar, Dominic Murphy, Justin Garson, Philip Gerrans, Rachel Cooper, Steeves Demazeux, Leen De Vreese, Maël Lemoine, Tim Thornton, Andreas De Block, Jonathan Sholl
Edited by: Luc Faucher, Denis Forest
DOI: 10.7551/mitpress/9949.001.0001
ISBN (electronic): 9780262362931
Publisher: The MIT Press
Published: 2021

The open access edition of this book was made possible by generous funding and support from Arcadia – a charitable fund of Lisbet Rausing and Peter Baldwin



The MIT Press

The open access edition of this book was made possible by generous funding from Arcadia—a charitable fund of Lisbet Rausing and Peter Baldwin.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Faucher, Luc, 1963– editor. | Forest, Denis, editor.
Title: Defining mental disorder : Jerome Wakefield and his critics / edited by Luc Faucher and Denis Forest.
Description: Cambridge, Massachusetts : The MIT Press, [2021] | Series: Philosophical psychopathology | Includes bibliographical references and index.
Identifiers: LCCN 2020016671 | ISBN 9780262045643 (hardcover)
Subjects: LCSH: Wakefield, Jerome C. | Psychiatry--Philosophy. | Mental illness--Philosophy. | Mental illness--Diagnosis. | Mental illness--Classification.
Classification: LCC RC437.5 .D434 2021 | DDC 616.89--dc23
LC record available at https://lccn.loc.gov/2020016671